

2004

High Stakes Testing in Louisiana: An Analysis of the Disparate Impact on Black and White Eighth Grade Students and the Perspectives of Parents

Monica L. Mancuso

Louisiana State University and Agricultural and Mechanical College

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_dissertations



Part of the [Education Commons](#)

Recommended Citation

Mancuso, Monica L., "High Stakes Testing in Louisiana: An Analysis of the Disparate Impact on Black and White Eighth Grade Students and the Perspectives of Parents" (2004). *LSU Doctoral Dissertations*. 486.
https://digitalcommons.lsu.edu/gradschool_dissertations/486

This Dissertation is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Doctoral Dissertations by an authorized graduate school editor of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.

**HIGH STAKES TESTING IN LOUISIANA: AN ANALYSIS OF THE
DISPARATE IMPACT ON BLACK AND WHITE EIGHTH GRADE STUDENTS
AND THE PERSPECTIVES OF PARENTS**

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
Requirements for the degree of
Doctor of Philosophy

in

The Department of Educational Leadership, Research,
and Counseling

by

Monica L. Mancuso

B.S., Louisiana State University, 1975

B.A., University of Louisiana at Lafayette, 1982

M.Ed., Nicholls State University, 1989

December 2004

DEDICATION

I dedicate a page in my dissertation to my dear family and friends;

To my husband, Bart, who encouraged and supported me in furthering my education,

To my children, Frances, Mary, Andrew, and Luke (and their spouses), who are my
treasures,

To my mother, Mary Catherine Asping Laughlin, who taught me persistence and the
value of an education,

To my brothers and sisters, Helen, Karen, Mary Jo, Earl, and Brian, who understand my
search for knowledge,

To my dear friend, Janice O. Verret, who accompanied me on road trips and to Su and
Rebecca who supported me,

To my friends, Barbara, Paula, Calley, Catherine, Mary, Marcia, Shirley, and Molly
who inspire me, and

To my classmates, Jackie, Dorian, Matt, Menthia, and Joe who shared their friendship in
the dissertation journey.

ACKNOWLEDGEMENTS

I am grateful for people who have supported me and provided guidance in my courses of study and in the dissertation process. The help I received improved my research.

I first wish to acknowledge my major professor, Dr. Dianne Taylor. You taught me how to reach higher goals and not to accept less than the best. You read and responded to my work with valuable recommendations. I appreciate your efforts in guiding my learning. I am grateful to Dr. Amy Dellinger for guiding me in the quantitative data analysis and for encouraging me through her words.

I also wish to acknowledge my dissertation committee. I am grateful to Dr. Spencer Maxcy for encouragement and guidance in getting my writing published. Dr. Charles Teddlie was helpful in focusing my research and I am grateful for his clear and timely comments. I wish to acknowledge Dr. Kim MacGregor for providing resources to guide my research and making useful suggestions. I appreciate Dr. Tim Page who served as the dean's representative. In all, I am humbled by the expertise of the members of my committee. The community of Louisiana State University is fortunate to have such talented professors among their ranks.

TABLE OF CONTENTS

DEDICATION.....	ii
ACKNOWLEDGEMENTS.....	iii
LIST OF TABLES.....	vii
LIST OF FIGURES.....	ix
ABSTRACT.....	x
CHAPTER 1. INTRODUCTION.....	1
Overview.....	2
Chaos Theory as a Lens for Examining Standards Based Reform.....	5
Standards Based Reform Development.....	7
School Accountability: The Louisiana Context.....	10
Purpose.....	14
Importance.....	17
Research Questions.....	18
Definition of Terms/Variables.....	19
Delimitations/Limitations.....	21
Organization.....	22
CHAPTER 2. REVIEW OF RELATED LITERATURE AND RESEARCH.....	24
Development of Standards Based Reform.....	24
Federal Involvement.....	25
Louisiana Response.....	32
Review of Research on Testing: Roles and Attributes.....	36
Support and Challenges.....	36
What and Who Gets Tested.....	39
Testing Impacts on Classroom Practices.....	43
Testing Impacts on Students.....	47
Appropriate Test Use.....	51
Related Research on Chaos Theory.....	54
Heisenberg Uncertainty Principle.....	54
Complexity Theory.....	55
Chaos Theory.....	56
Summary.....	61
CHAPTER 3. METHODOLOGY.....	63
Phase I.....	63
Phase II.....	66
Sampling.....	68
Instrumentation.....	70
Data Collection and Analysis.....	71
Trustworthiness.....	72

Summary.....	73
CHAPTER 4. QUANTITATIVE RESULTS.....	74
Phase I.....	74
Number and Location of Students.....	75
Race and Community Type.....	77
Pass/Fail Categories.....	81
Achievement Levels.....	83
Magnet School Impact.....	85
Summary.....	87
CHAPTER 5. QUALITATIVE RESULTS.....	89
Sampling.....	89
Focus Group Parent Reports.....	92
Country Road, Black, Ninth Grade.....	96
Why Were Students Tested?.....	96
What Were Students Learning?.....	96
Was the Test Fair?.....	97
City Street, White, Ninth Grade.....	98
Why Were Students Tested?.....	98
What Were Students Learning?.....	99
Was the Test Fair?.....	100
City Street, Black, Grade 8.5.....	102
Why Were Students Tested?.....	103
What Were Students Learning?.....	103
Was the Test Fair?.....	105
City Street, White, Grade 8.5.....	107
Why Were Students Tested?.....	108
What Were Students Learning?.....	108
Was the Test Fair?.....	109
Analysis By Themes.....	111
Theme 1 – Reform Intent.....	111
Theme 2 – Student Learning.....	115
Theme 3 – Test Fairness.....	119
Parent Perspectives on the Disparate Impact.....	120
Testing for Public/Non-public School Students.....	124
Single Criterion.....	128
Timing of LEAP 21.....	129
Summary.....	134
CHAPTER 6. DISCUSSION, CONTRIBUTIONS, AND RECOMMENDATIONS...137	
Discussion.....	137
Policy Considerations and Recommendations.....	142
Summary.....	148

REFERENCES.....150

APPENDIX A: LETTERS – SUPERINTENDENT, PRINCIPAL, AND PARENT.....166

APPENDIX B: FOCUS GROUP QUESTIONS.....169

APPENDIX C: FORM FOR POST FOCUS GROUP NOTES.....171

VITA.....172

LIST OF TABLES

Table 1. Descriptors of the Substance of SBR Examined Through a Chaos Theory Lens..	7
Table 2. How Community Types are Defined.....	20
Table 3. LEAP 21 Policy Requirements and Placements for Taking the Grade 8 Test ...	35
Table 4. Negative High Stakes Test Impacts on Students.....	48
Table 5. Parent Focus Group Sampling.....	69
Table 6. Eighth Graders by Race and Community Type.....	75
Table 7. Multivariate Analysis of Variance (MANOVA) Results.....	77
Table 8. Mean Scaled Scores by Race and Community Type.....	78
Table 9. <i>t</i> -tests and Effect Sizes for the Interaction of Race and Community Type.....	80
Table 10. Mean ELA and Mathematics Scores Across Community Type.....	80
Table 11. Mean ELA and Mathematics Scores by Race.....	81
Table 12. Passage Rates in ELA Across Race and Community Types.....	82
Table 13. Passage Rates in Mathematics Across Race and Community Type.....	83
Table 14. Achievement Levels for Grade 8 LEAP 21.....	83
Table 15. ELA Achievement Levels by Race and Community Type.....	84
Table 16. Mathematics Achievement Levels by Race and Community Type.....	85
Table 17. Passage Rates With and Without Magnet School Student Scores.....	86
Table 18. Focus Group Sampling.....	90
Table 19 Four Focus Group Reports.....	93
Table 20. Test Purpose.....	112
Table 21. Educational Improvement.....	116
Table 22. Test Disparities for Groups of Students.....	122

Table 23. Disparities of Public and Non-public School Test Requirements.....125

Table 24. When LEAP 21 Testing Occurs.....130

LIST OF FIGURES

Figure 1. How Tracing SBR Is Used to Understand the Study.....	8
Figure 2. Louisiana District Map Showing Schools by Location.....	76
Figure 3. ELA Mean Scaled Scores by Race and Community Type.....	78
Figure 4. Mathematics Mean Scaled Scores by Race and Community Type.....	79

ABSTRACT

Standards based reform (SBR) measured by LEAP for the 21st Century (LEAP 21), the high stakes testing program in Louisiana, was explored across rural, suburban, and urban school community types. Differences in scores of Black and White eighth graders on LEAP 21 were analyzed using a multivariate analysis of variance (MANOVA), revealing race and the interaction of race and community type to be statistically significant at $p < .05$. Black students did not score as well as White students when scores were analyzed by pass/fail and achievement level differences. The percentage of Black students passing the test increased at rates greater than White students when scores from magnet schools were included.

Parents of eighth graders disagreed with the use of a single criterion for promotion. Parents said the instruction students received, pace of instruction, and stress of testing negatively affected student performance with the result that their children considered a Graduate Equivalency Diploma (GED) or quitting school. Parents agreed that affluence and the level of family support affected students' test performance. Some parents also believed it was unfair that non-public school students were not required to take or pass such tests and some parents considered enrolling their child in a non-public school to avoid the consequences of the state testing program.

CHAPTER 1. INTRODUCTION

Across the country, students who are exiting high school without requisite knowledge and skills are costly to society. Least costly are college-bound students who need remedial classes to prepare them for college course work (NCES, 1996; Ravitch, 2000) and students requiring workplace training for low-skill jobs (Jaeger, 1982; Ravitch, 2000). Most costly are students whose lack of basic skills combines with unacceptable social behavior, leading them to involvement in illegal activity and ultimately incarceration (Cardenas, 1994). At the same time, public policy focus in education is on holding schools, teachers, and students accountable for student learning.

In developing accountability programs, states identified acceptable levels of knowledge across subject areas to be attained by exiting students and codified the levels in educational standards and benchmark knowledge, a part of standards based reform (SBR). Students are expected to attain proficiency in benchmark knowledge and demonstrate the proficiency by taking tests designed to measure the content standards (Rothman, Slattery, Vranek, & Resnick, 2002). State tests anchored in state developed content standards are administered in certain gatekeeper grades, including high school grades. In Louisiana, the results of the tests determine whether fourth and eighth graders pass to the next grade and, at the high school level, whether students graduate. As a result of these consequences, these tests have come to be called high stakes tests. Thus, a student with passing course grades who fails to meet the state-determined level of proficiency fails to move on in school or fails to graduate.

Testing provides feedback for a variety of purposes including identifying student strengths and weaknesses, providing evidence that students learn, and providing public

accountability (Lewis, 2000). Each of the uses benefits the education of American youth. However, a historical examination indicates that test scores have been used to keep African-American and Latino students in segregated schools (Chachkin, 1989; Garcia & Pearson, 1993). More recently, testing has been used to send minority students disproportionately into special education (Chachkin, 1989; Garcia & Pearson, 1993) and lower tracks in middle and high schools (Gamoran, 1996; Garcia, Stephens, Koenke, Pearson, Harris, Jimenez, 1989; Rebell, 1989). Consequences for these students have devastating, life-long effects; the consequences for society are not incidental and can be costly as noted above.

Because of prior misuses, testing with high stakes consequences merits a high level of scrutiny. The scrutiny needs to go beyond longitudinal improvements in passage rates to include monitoring the impact on groups of students who have been previously disenfranchised in the educational process—financially poor and minority students.

Overview

The impact of testing on students will be further developed in Chapter Two; however, a study by Johnson and Johnson (2002), reported in *High Stakes: Children, Testing, and Failure in American Schools*, suggests that poor and minority students in Louisiana are impacted disparately because of high stakes testing. The impact is disparate because it suggests minority students of the state have not received the needed instructional program to pass the prescribed test (Popham, 2000), but that more affluent students have received such a program. To support this claim, Johnson and Johnson spent a year as classroom teachers in a rural Louisiana school, preparing students for standardized testing. Their research revealed the lack of a facility to accommodate a learning climate, an absence of

meaningful teaching materials, the lack of appropriate professional development, and the ambiguity of interpreting content standards to prepare students for what might be covered on the test. The findings in the Johnson and Johnson (2002) study clearly suggest two areas of concern. The first is a lack of alignment between what should be taught and what will be assessed. The second is a lack of capacity evidenced by inadequate school facilities, instructional materials, and professional development. In effect, the intent of having all students meet high academic standards fell short because of the lack of alignment and capacity. Credibility in the Johnson and Johnson study was enhanced by the fact that both researchers had taken a leave of absence from their university teaching positions to conduct their research. At the Louisiana university, the researchers prepared students to become teachers, in effect; they prepared others to work in standards based instructional classrooms. In spite of the level of knowledge of the researchers, a disorderliness or chaos existed about the broad benchmarks and what would be assessed. The delivery of instruction, compromised by lack of adequate school facilities and instructional materials, did not effectively meet learner needs.

The researchers contrasted the impact of high stakes, criterion referenced tests (CRTs), known in Louisiana as the Louisiana Educational Assessment Program (LEAP) and more specifically as *LEAP for the 21st Century* (LEAP 21), between their largely poor, minority students and their affluent, mostly White counterparts attending nonpublic schools in the state. The affluent White students:

are sent to the private Deerborne Academy. . . .The annual tuition rivals that of a small university, \$3,000-\$4,000 per child per yearA benefit of attending Deerborne or any of the numerous private schools in Louisiana is that the children do not have to take the LEAP test and subject themselves to anxiety over the test and possible repetition of a grade. We have learned that some families are willing to pay the tuition

just to avoid the consequences of high stakes testing mandated by the state for the public schools (p. 20).

Children of affluent parents in Louisiana may avoid meeting challenging academic standards by choosing to attend a nonpublic school. For the children of less affluent parents, whose only choice is the local public school; the testing component of standards based reform (SBR) can mean grade failure. Parent perspectives on using nonpublic schools as a means to avoid these consequences of high stakes testing merits further exploration.

To fully explore the disparate impact of high stakes testing, the current study described the evolution of SBR, analyzed statistical differences of test scores between Black and White students, and explored parent perspectives of testing impacts on their eighth and ninth grade children. The current study used part of the review of literature to understand SBR development, presenting the range of SBR stakeholders and illustrating the reform development process that resulted in high stakes testing. The review of SBR development also allowed understanding of current conditions as influenced by the past. The current study examined differences in performance on the Louisiana high stakes, eighth grade test between Black and White students across the state by use of statistical analyses. The examination provided a closer look at the challenge of the accountability program in meeting student learning needs in a state with a high minority population (U.S. Census Bureau, 2003a). In addition, interviews that explored the perspectives of parents were conducted. Parents of students who both passed and failed the test were included and were asked about the test and its impact on their child's learning. The overview of this study provides insight into how the study evolved; however, the lens by which SBR may be judged situates the study for the reader.

Chaos Theory as a Lens for Examining Standards Based Reform

Chaos theory provided an aid for studying the disparate impact of high stakes testing. In use of chaos theory, critic Maxcy (1995) described the substance of the studied phenomenon as “multi-leveled, complex, recursive, dissipative, and above all disorderly” (p.39). The studied phenomenon is a) multi-leveled because it must be viewed as a system and more than a sum of the parts (Gleick, 1987; Maxcy, 1995), b) complex because transitions are not smooth; c) recursive as order is transitional, d) dissipative as the initial circumstances take “unpredictable paths” (Maxcy, 1995, p.37), and d) disorderly as the phenomenon lacks one single goal. Here the substance is the organization, dynamics, and language of SBR instructional delivery, of which high stakes testing in Louisiana is a part.

Chaos theory applies as a lens to view SBR from a distance. SBR is multi-leveled as it involves federal, state, and local governmental roles and the roles of each level have shifted away from local control of learning in schools, as will be developed more thoroughly in Chapter Two. However, with the shift in governmental roles affecting education, the interactions of the levels of government became more complex, which is also a descriptor of the substance of chaos theory. Previously, local boards controlled schools with minimal state or federal direction for instructional delivery (Fowler, 2000). However, schools responded to the shift in governmental roles ultimately changing the manner in which instruction is delivered and learning takes place (Fowler, 2000). In addition, other stakeholders such as professional organizations, teachers, and parents are involved in SBR, adding to the complexity. The substance of the phenomenon examined through the lens of chaos theory fits SBR as being recursive and dissipative in that the

focus of education changes. While tenets of equality from prior reforms remain, excellence, accountability, and choice have become central (Boyd & Kerchner, 1988; Fowler, 2000). By calling for disaggregated data in *No Child Left Behind* (NCLB), equality and equity are becoming the focus again (Gordon & Bonilla-Bowman, 1994), making SBR a part of a recursive and dissipative cycle.

Especially important in the current study was the sociopolitical context conceptualized by the arbitrariness of cutoff scores (Ellwein & Glass, 1989; Heubert & Hauser, 1999), adding to the disorderliness of SBR. For example, consider the student who passes all coursework and is recommended by the teacher for passing to the next grade; however, the student does not meet a cutoff score on the high stakes test. The school organization allows the student to experience mixed feedback on performance as meeting the cutoff score is used as a measure for school accountability and student accountability. In effect, SBR implementation becomes disorderly. Table 1 provides a description of how the substance of SBR is examined through the lens of chaos theory.

Standards based reform involves a predetermined state of reality. Situated within a sociopolitical context, the state of reality reflects an acceptable passage rate for the test. Not everyone is allowed to pass, lest the test be criticized for being too easy (Meier, 2002). Neither can the passage rate be set at a level where few students pass. Therefore, predetermined politically acceptable rates of passing and failing (Ellwein & Glass, 1989; Heubert & Hauser, 1999) introduce chaos to testing programs that purport to expect high academic standards (Linn, 2000; NCLB, 2001). The current study uses chaos theory as the lens through which to examine the disparate impact of high stakes tests on poor and

minority students. Use of part of the literature review to understand SBR development is described in the next section.

Table 1.

Descriptors of the Substance of SBR Examined Through a Chaos Theory Lens

Description	Evidence
Multi-leveled	Federal level calls for disaggregation of data; the state level sets standards; the local level delivers standards based instruction.
Complex	A shift in roles of government took place; many stakeholders with varying levels of understanding are involved.
Dissipative	Equality dissipates to accountability, excellence, and choice.
Recursive	Equality returns in disaggregation of data.
Disorderly	Mixed feedback is given to students on performance, confusion occurs about the standards and what is tested.

Note. Descriptions for SBR substance examined through chaos theory come from Maxcy, (1995, p.39).

Standards Based Reform Development

The present study traced the development of standards and assessment within the evolution of educational reform. Reform occurs in response to prior research findings. More importantly, reform occurs in response to policies and politics at state and federal levels (Smith, Miller-Kahn, Heinecke & Jarvis, 2004). Thus, reform in education is not based solely on best practices supported by research; but, it is situated within a changing political and societal context. The tracing of SBR development and a lens of chaos theory in the current study allowed a thorough examination of SBR in the context of

instructional practices and sociopolitical decision making guiding many reform processes.

Figure 1 depicts the development of SBR situated in a sociopolitical context with the aim of high academic achievement for all students (Linn, 2000; No Child Left Behind, 2001); but, the predetermined reality is politically acceptable passage rates. In addition, Figure 1 portrays SBR as incorporating two main concepts, curriculum alignment and capacity building (O'Day & Smith, 1993; Smith, O'Day, & Cohen, 1990). Both concepts

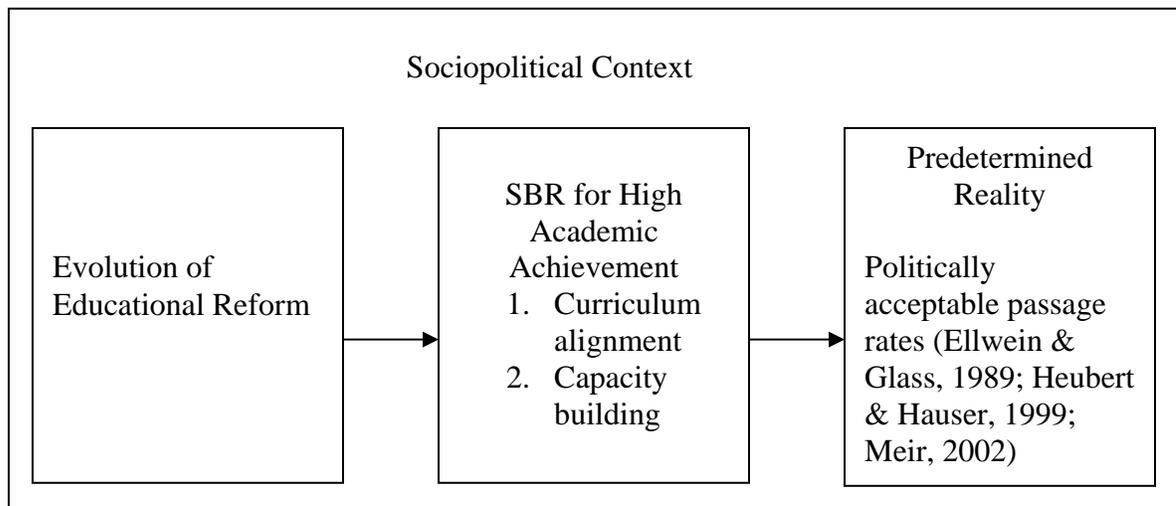


Figure 1.

How Tracing SBR is Used to Understand the Study

are important in understanding the disparate effect of high stakes testing. Curriculum alignment involves matching the curriculum to standards and using testing to assess student performance on the standards. The quality of the instruction that takes place is absent from the discussion. Therefore, what is measured is how well the standards and assessment are aligned without discussion of how well the curriculum is taught (Airasian & Madaus, 1983; Le & Klein, 2002) or learning takes place. Capacity building involves

having teachers know what to teach with the organization acting to support their efforts (Carnoy, & Loeb 2002). Important to the current study is that students are held to high stakes consequences, but not all teachers are familiar with standards and benchmarks that are assessed (Johnson & Johnson, 2002). Both a lack of clearly aligned standards and teacher capacity hinder SBR implementation and because of their importance, the two concepts are discussed more thoroughly in Chapter Two. Suffice it to say here that knowledge of past events provides a background for understanding current educational reforms (Donato & Lazerson, 2000; San Miguel, 1987; Walker, 1996).

While many reasons exist to present SBR development, the current study used the reason of informing the way we judge school reform today (Gall, Borg, & Gall, 2003). The current study traced SBR development and provided the advantage of informing practice when policy decisions must be made (Gall, et al.2003). For example, the study presents information relevant to informing the practice of school improvement currently taking place in Louisiana schools. One of the practices addresses school improvement efforts for subgroups of students. In summary, tracing SBR development informs practice (Gall, et al, 2003; Graham, 1980).

Some researchers use historical information to provide insights and policy direction (Donato & Lazerson, 2000; Ravitch, 1995, 2000; Ravitch & Vinovskis, 1995; Vinovskis, 1999). For example, Ravitch (2000) provided insight into attempts in the early twentieth century to find alternatives to an academic curriculum for all students. In other words, policies were implemented using the reform of meeting individualized needs for noncollege bound students. The effect was a disparate one on “mainly the poor, immigrants and racial minorities-who were pushed into undemanding vocational,

industrial, or general programs by bureaucrats and guidance counselors who thought they were incapable of learning much more” (Ravitch, 2000, p.15). The current study took a similar approach, seeking to trace SBR development as a part of the literature review to provide insights into SBR policies in Louisiana. SBR development was traced both nationally and on the state level. In fact, the Louisiana context presents unique characteristics, which distinguishes the state from the rest of the nation. A description of the context of the study takes place in the next section.

School Accountability: The Louisiana Context

A description of the Louisiana context provides evidence that the state has one of the strongest accountability programs in the country; the accountability program includes high stakes tests for graduation and promotional purposes. Louisiana also has a high percentage of Black students attending schools in predominantly Black communities (Ravitch, 2000). Therefore, the state is a suitable context for studying the disparate impact of high stakes testing.

Carnoy and Loeb (2002) assigned Louisiana an accountability index of three on a scale of 0-6. The index ranked state testing systems based on consequences for schools and for students related to passing high school exit exams. Thirty-six states had scores of 0-2.5, which represented a weaker accountability index than the Louisiana index. Similarly, Amrein and Berliner (2002) used the number of consequences of testing to rank the strength of accountability programs in states. Accountability policies in Louisiana specify seven consequences including failure to graduate, retention in a grade, publication of school report cards, identification of low performing schools, closure of low performing schools, replacement of school personnel due to low test scores, and

enrollment of students from failing schools in other schools. Only Texas had more consequences tied to assessment (Amrein & Berliner, 2002). However, with the passage of a constitutional amendment in the fall of 2003 allowing the state of Louisiana to takeover low performing schools, Louisiana may lead the nation in tying consequences to testing in its schools. Additional support for the strong accountability program in Louisiana comes from data collected for the 2002-03 school year for Quality Counts (2003), using a survey of the 50 states and District of Columbia for reporting the progress of each state's educational system. Louisiana is one of five states that requires students to pass a test for promotional purposes and 1 of 19 states that ties graduation to passing a test (Education Week, 2004). Thus, high stakes testing for graduation is rather common in the nation, but testing for a promotional purpose is limited to a few states including Louisiana.

In Louisiana, minorities make up 36% of the population with Black being the largest minority group at 33%. The minority population of the U.S. makes up 25% of the population (U.S. Census Bureau, 2003a). In addition, in Louisiana 22% of the families with children 18 years and under fall below the poverty level (U.S. Census Bureau, 2003b). Simply stated, one in three people in Louisiana are classified minority and one in five families with young children live in poverty. Carnoy and Loeb (2002) found that states with a higher proportion of minority students implemented stronger accountability programs. Black student enrollment in Louisiana is 48%, making the public school system of the state second to Mississippi in concentration of Black students (Hoffman, Llagas, & Snyder 2003). The finding suggests that being held accountable is directed to

areas having high proportions of minority students of which Louisiana is one (Reardon, 1996).

Regional differences influence reforms. Elazar (1984) lists Louisiana as a part of the traditionalistic, political culture dominant in the South. Other regions, characterized as moralistic or individualistic cultures, respond to reforms within their own political culture. While other parts of the country are not immune to disparate impacts on minorities, the traditionalistic political culture is itself characterized by a resistance to change. SBR has meant setting high academic standards for all students (Linn, 2000) in public schools. Black students are disproportionately represented in public schools and White students almost exclusively comprise non-public schools. Moreover, as Johnson and Johnson (2002) reported, resources needed to facilitate students passing the test have not been made available to many public schools. Thus, SBR has not challenged the traditional, political culture in Louisiana.

Louisiana schools that score at the bottom in the accountability program are located in high-poverty areas including inner-city New Orleans, along the Mississippi River delta, and other pockets of poverty (Johnson & Johnson, 2002). The pattern of high-poverty reflects the outward migration, which took place between 1940 and 1960 in which 3 million Blacks left the South and moved north. The movement dropped the percentage of Blacks living in rural areas from 35 percent to 8 percent and the proportion of Blacks living in urban areas grew from 49 percent to 73 percent (Ravitch, 2000; Thernstrom & Thernstrom, 1997). Forty to 60 years later, the same areas in Louisiana reflect high poverty and low test scores (Johnson & Johnson, 2002).

Outward migration from the cities to the suburbs by Whites in the 1950s and 1960s added to high-poverty concentrations in Louisiana (Ravitch, 2000). Schools scoring at the top in accountability in Louisiana have selective admissions procedures, such as magnet schools and university laboratory schools or they are located in affluent suburbs (Johnson & Johnson, 2002). Community type, a variable in the current study, recognized the high and low poverty pattern in school locations.

In addition to community type, race is a variable in the current study. Studying the impact of high stakes testing on Black students is especially suited to Louisiana because of its strong accountability program (Amrein and Berliner, 2002; Carnoy & Loeb, 2002; Education Week, 2004), high minority population (U.S. Census Bureau, 2003a), and socio-economic pattern that breaks along racial lines. Previous research suggests that Black students do not test as well as White students (Applebee, Langer & Mullis, 1987; Educational Testing Service, 1988; Garcia & Pearson, 1993; Mullis & Jenkins, 1990; National Center for Education Statistics, 1988). Thus, a high poverty level coupled with a high minority population in a state with a strong accountability program may support a continuation of conditions described earlier in the history of Louisiana. Johnson (1942) described school conditions for Black students in Louisiana as “a vicious circle” (p.109). During the 1940s, Black children in Louisiana left home from a deprived setting to attend a school that reflected the same deprivation. The schools for the children were comprised of dilapidated buildings and meager equipment. Schools in Louisiana became part of the circle of deprivation for Black children (Johnson, 1942; Ravitch, 2000). In effect, schools maintained the same deprived conditions of home. Similar deprived conditions exist in

some Louisiana schools for Black students held accountable through high stakes tests (Johnson & Johnson, 2002).

In addition, Louisiana has a large nonpublic school population including private and parochial students. Sixteen percent of the Louisiana school population attends nonpublic schools compared to 10% nationally (LDE, 1999). Moreover, nonpublic school students are not mandated to take or pass the state test for graduation or promotion purposes. Simply, one in six students in Louisiana is not required to take a high stakes test but is still promoted to high school and exits high schools receiving a diploma that carries the same privileges as the one issued by public schools (LDE, 2000b). Students who are not required to take the high stakes tests are those whose parents can afford tuition at a nonpublic school (Johnson & Johnson, 2002). Therefore, the present study uses access to nonpublic schools as a component of the disparate impact the accountability program places on Black students in the state. Especially relevant to the Louisiana context is the high poverty of the state compared to the rest of the nation. In a listing of states by rank of median household income, Louisiana ranked 49th with a median household income of \$33,311 (U.S. Census Bureau, 2002), making nonpublic school tuition not a consideration for most poor families.

Purpose

The fact that Black students do not test as well as White students on formal tests (Applebee, Langer & Mullis, 1987; Educational Testing Service, 1988; Garcia & Pearson, 1993; Mullis & Jenkins, 1990; National Center for Education Statistics, 1988) nor on high stakes assessments (Gladfelter, 2000; Lewis, 2000) is documented in the literature. Recognizing a potential disparate impact on racial and ethnic minorities, a

provision of the NCLB legislation of 2001 calls for disaggregated data in testing according to race, among other criteria, to monitor how well the achievement difference is being closed for groups of students. Criticism surrounds the disaggregated approach because distortion occurs in oversimplifying the achievement differences to a comparison of scores on a test (Darling-Hammond, 1992; Dorn, 1998; Koretz, 1992; Koretz & Diebert 1993; Shepard, 1991). A phase of the current study examined the statistical significance of the difference in scores of Black students compared to scores of White students by examining the magnitude of the test score differences on LEAP 21 across all community types.

The current study also examined the impact of testing beyond test scores through use of parent focus group interviews. Focus groups were used to gain parent perspectives in areas of meeting the reform intent by raising standards in our schools, increasing student learning, and impacting all groups of students similarly. Findings from this part of the study suggest that the tests lack the full reform intent of raising student achievement, as Black students lack quality instruction and resources for passing the test. Also, parent perspectives suggest that Black students are additionally affected by timing the test right before entry into high school and right as struggling students become old enough to dropout of school. Lastly, parent perspectives suggest that basing promotion on a single test for students in public schools and exempting nonpublic schools from testing is unfair and creates a dual system due to test requirements.

Tracing SBR development as a part of the literature review allowed present conditions of SBR, particularly high stakes testing, to be revealed by using the past (Gall et al., 2003). Louisiana offered a useful sample because the state has a (a) large, Black, high

poverty population and a significant, White, more affluent population, and (b) public school system as well as a high number of nonpublic schools, including religious and non-sectarian. The large number of nonpublic schools is important because it offers an option to families that can afford tuition and thereby side-step the state accountability law, including high stakes testing.

The present study explored the effects of high stakes testing across various contexts. The context included the type of community where students attend school; rural, suburban, or urban. Two questions guiding the study asked: First, what is the extent of variation in high stakes test scores between Black and White eighth graders across community types? And, second, do parent perspectives of the Louisiana accountability high stakes testing program vary based on the child being Black or White, having passed or failed the test, and attending a rural, suburban or urban school?

The current study used scores from eighth graders because the potential for the student dropping out increases for those not passing the test by denying access to the regular high school curriculum. The students are only eligible to earn an alternative Graduate Equivalency Diploma (GED) or vocational program certificate (LDE, 2001a). The students also approach the age of 16, whereby, students are old enough to drop out of school with parental permission. Thus use of eighth grade scores helps to examine the full disparate impact of high stakes tests on students. While understanding the purpose of the present study is relevant, discussion also takes place on the importance of the study leading beyond what the literature already supports. A discussion of the importance of the study takes place in the next section.

Importance

While prior research supports a gap in achievement between Black and White students, the current study described findings measuring statistical significance of the difference in student scores across school community contexts and also explored parent perspectives of high stakes testing for their child. Previous studies examined teacher perspectives of high stakes testing and the impact of testing on the curriculum (Center for the Study of Testing, Evaluation, and Educational Policy, 1992; Garcia, & Pearson, 1993; Rothman, 1992). Few studies have examined parent perspectives (Hamilton & Stecher, 2002). The current study addressed the need.

In addition, research was missing describing parent perspectives based on whether their child passed the assessment and whether families had access to nonpublic schools. Moreover, previous studies failed to describe parent perspectives of Black families denied options because of poverty and race. Therefore, the current study contributed to the literature, filling in gaps in understanding the impact of high stakes testing. The current study further explored a reductionist strategy simplifying the goals and practices of schools to an arbitrary score on a single test. The question becomes not what learning takes place by the student, but did the student meet the score to pass. The study explored this narrow manner of measuring student achievement. Also, the current study suggests a need for changes in the high stakes testing policy of Louisiana whereby less of a disparate effect on Black students takes place. For example, consideration of more opportunities to pass grade 8 and multiple measures for promotion purposes are revealed through parent interviews. The present study was guided by research questions that

explored the disparate impact of high stakes testing on Black students of Louisiana. The next section presents the research questions.

Research Questions

To explore the disparate effect of high stakes testing in the state of Louisiana, the study used the following research questions:

1. What were the forces and events leading to high stakes testing for promotion purposes in Louisiana?
2. What are the differences in LEAP 21 ELA and mathematics scaled scores between Black and White eighth graders in rural, suburban, and urban community types?
3. What are the differences in LEAP 21 ELA and mathematics pass/fail rates between Black and White eighth graders in rural, suburban, and urban community types?
4. What are the differences in LEAP 21 ELA and mathematics achievement levels between Black and White eighth graders in rural, suburban, and urban community types?
5. What are the differences in LEAP 21 ELA and mathematics passage rates between Black and White eighth graders in rural, suburban, and urban community types when magnet school students are included versus when they are not?
6. What are the perspectives of parents of eighth graders about why students take LEAP 21? Parent perspectives were explored for both meeting the reform

intent of raising standards in schools and increasing student learning by use of high stakes consequences.

7. What are the perspectives of parents of eighth graders about the fairness of the test on groups of students? Test fairness was explored for impacting groups of students who are rich or poor, Black or White, male or female, and public or non-public.
8. What are the perspectives of parents of eighth graders about using the test as a single criterion for promotion and timing the test as students complete grade 8?
9. Do perspectives of parents of eighth graders who pass and fail LEAP 21 vary based on race, Black and White, and community type; rural, suburban, and urban?
10. Do perspectives of parents of eighth graders who take LEAP 21 suggest a disparate impact on poor and minority students in Louisiana?

To fully understand the questions, terms and variables as they are used must be given.

Defining the terms of the study takes place in the next section.

Definition of Terms/Variables

To clarify understanding specific to the present study, this section provides meanings for terms and variables. SBR is defined as reform efforts in which teachers teach to Louisiana state standards and hold students accountable through performance based assessment. Performance based assessment at grade 8 is a criterion referenced test (CRT) for promotional purposes, LEAP 21. Success on the high stakes test for promotion is defined as students who meet or exceed the predefined cutoff score on the English

Language Arts (ELA) and mathematics subtests. Benchmark knowledge is defined as levels of performance particular to a grade at which a student is expected to have attained proficiency. Disparate impact is defined as having the effect of creating significant differences for groups of students (Heubert & Hauser, 1999).

Community type definitions follow common core of data reporting for schools, which take place annually. See Table 2.

Table 2

How Community Types are Defined

The Current Study	Louisiana Department of Education (LDE)
Rural	Small town
	Rural outside of a Metropolitan Statistical Area (MSA)
	Rural inside of a MSA
Suburban	Urban fringe of a large city
	Urban fringe of a mid-size city
	Large town
Urban	Large city
	Mid-size city

Note: Community types used by the Louisiana Department of Education (LDE) come from the 1990 U. S. Census.

School systems nationwide report data on student enrollment. The schools are categorized according to eight school-community types based on the physical or mailing address of the school and population density (LCES, 2002). The current study coded the eight school community types into three broad areas; rural, suburban, and urban. The

U.S. Census Bureau (USDOE, 2002) uses similar categorizations. For purposes of this study, community type refers to the physical location of the school. Rural means small town, rural outside of a Metropolitan Statistical Area (MSA), or rural inside of a MSA. Suburban means urban fringe of a large or mid-size city, or a large town. Urban means a large or mid-size city. Thus the current study used three common community types; rural, suburban, and urban, to describe school locations in Louisiana

Definitions of terms specific to a study are important and have been provided in this section; however, delimitations and limitations of a study also provide important information and are given in the next section

Delimitations and Limitations

Delimitations of the present study included the use of test scores for eighth graders and not fourth graders. Also, the scores of interest are those from CRT testing and not norm referenced testing (NRT). A small percentage of minority students in Louisiana (NCES, 2003) are Hispanics (1%), Asians (1%), and American Indians (<1%) hence, the study is delimited to Black students.

The study is delimited to public schools that are not alternative, charter, or university lab schools. Most commonly, alternative schools restrict entry to students with discipline referrals; charter schools involve parent choice; and university lab schools are affiliated with a university and are few in number. The schools do not lend themselves to study inclusion. Because the study uses community type as a variable, magnet schools could not be excluded; however, the impact of the schools is also explored. Twenty-six magnet schools were included in the study with 21 urban, 4 suburban, and 1 rural school testing eighth graders in the state.

The current study used scores as absolutes, and therefore the arbitrariness of cutoff scores and inferences made by them limit the study. For example, inferences of student achievement reflect performance as measured by the assignment of a single score and not other indicators of student performance. The Louisiana context limits the study. Because Louisiana presents a unique context, results may not generalize to other states. The study is also limited by the volunteer nature of the sample. Parents who declined to participate in focus group interviews may hold different opinions from parents who agreed to participate. The extent to which opinions vary is not known.

Organization

Organization of the dissertation follows a format in which Chapter One presents a concise description of the study, including the conceptual framework, research questions, term definitions, delimitations and limitations, and organization of the dissertation.

Chapter Two provides a review of related literature and research on educational reform. In the review, tracing the involvement of federal, state, and local stakeholders in the reform process takes place. In addition, issues regarding high stakes testing implementation in Louisiana are discussed. The chapter concludes with a discussion of chaos theory and its application to the reform process.

Chapter Three presents the methodology used in the study. The methodology involved two phases. Phase I explored differences in test performance between Black and White students across three community types. Phase II involved interviews exploring parent perspectives regarding the impact of high stakes testing on their children. Instrumentation, data collection, and data analyses are described. The chapter concludes with presentation of methods for ensuring the trustworthiness of the results.

Chapter Four provides the quantitative results of the study. A quantitative description of test score differences is provided. The description provides a comparison of Black and White student differences across the three community types by mean scaled scores, pass/fail rates, and achievement levels. A multivariate analysis of variance (MANOVA) is used to determine statistical significance of test score differences in ELA and mathematics. Post hoc *t* test results are given to examine specific differences with Cohen's *d* used to give effect sizes.

Chapter Five presents the qualitative results of the study. Focus group data across community types of parents of Black and White students who both passed and failed the test were examined. Four reports of focus group sessions were given to allow the reader into the conversation of the parents. The data were then examined for differences in perspectives by themes and by categories of parents.

Chapter Six presents discussions, contributions, considerations, and recommendations. The findings are discussed in relation to the manner the study contributes to the literature. Policy considerations are then made as suggested by the findings. The chapter concludes with recommendations for further study.

CHAPTER 2. REVIEW OF RELATED LITERATURE AND RESEARCH

The current study examined the difference between test scores of Black and White Louisiana eighth grade students, and perspectives of parents of eighth graders regarding the Louisiana accountability high stakes testing program. To situate the study in the literature for the reader, several areas of research were reviewed and are presented in this chapter.

The role of federal, state, and local governments is given. The section provides a historical chronology of the development of standards based classrooms and student assessment providing background for the current study. Next, the literature review shifts to focus on testing expansion, a part of standards based reform (SBR). Discussion takes place about what content will be tested and the impact of testing on classroom practices and students. A contrast of the Louisiana test instrument is made against other measures of appropriate test use. The contrast provides a review of the quality of the Louisiana test as a measure with high stakes consequences for students. Finally, chaos theory is examined to provide a lens through which to examine SBR implementation, including high stakes testing. Also, chaos theory is contrasted with Heisenberg's Uncertainty Principle and complexity theory as used by O'Day (2002) to reveal how the use of other theories compares as a lens for SBR. The next section begins with federal initiatives involving SBR and traces the reform to Louisiana.

Development of Standards Based Reform

A description of federal government involvement, which led to changes in educational control, takes place in this section. The chronology of events led to reforms ultimately affecting the manner students are currently taught and held accountable (Fowler, 2000).

When states received more discretion in federal fund spending, states took a greater role in how local educational implementation would take place (Fowler, 2000; Kaplan & O'Brien, 1991). The implementation involved meeting state content standards measured through statewide assessment, a function previously centered in federal monitoring of federal programs in schools attended by low-income students. Also included in this section is the response of Louisiana to the federal policy changes. In effect, the reform response impacted students in Louisiana through increased state mandated testing.

Federal Involvement

The discussion of federal involvement in SBR includes factors which increased federal involvement, calls for reform, including testing expansion, and commitment from both the federal and state executive branches. Federal involvement in education has historically been less prevalent, as states take their responsibility for education from the Tenth Amendment of the U. S. Constitution. Specifically, powers not delegated to the federal government nor prohibited by it are reserved to states. Education is one of these powers; however, federal influences exist as will be shown throughout this section. The call for standards began in the 1950s when President Dwight Eisenhower responded to public demand for higher standards and greater attention to mathematics, science, and foreign languages being taught in the schools (Ravitch, 2000). Support resulted from perceived poor quality of education taking place in U.S. schools at the time when the Soviet Union *Sputnik* space satellite launching took place. The response included passage of the National Defense Education Act (NDEA) providing increased funding in K-12 and higher education to meet the academic challenge made by the satellite launching..

Social upheavals of the 1960s and 1970s, including the civil rights struggle and anti-war demonstrations, led President Lyndon Johnson to increase federal spending in education. The increased funding was a part of the War on Poverty of the Johnson administration in which education became “a part of a larger struggle for social, political, and economic equality” (Timar, 1994, p.67). The federal Elementary and Secondary Education Act (ESEA) of 1965 passed Congress and granted compensatory educational services to disadvantaged students through its Title I program assisting local educational agencies to serve large concentrations of low-income families (ESEA, 1965; Smith & Jenkins, 1982).

Federal reform initiatives in high stakes testing began to take shape in the 1980s. Federal educational policy toward K-12 schools changed to the New Federalism under President Reagan (Fowler, 2000). State governments had not been very involved in public education; therefore, “state governments delegated most of their authority over public education to local school districts without requiring them to do much to demonstrate accountability” (Fowler, 2000, p.3). Under the Reagan administration, federal educational spending for U.S. schools declined from 8% to 6% (Fowler, 2000; Stroufe, 1995) of the national budget and Title I funding that had increased by 350% from 1966 to 1986 (Natriello & McDill, 1999) was also subjected to decreased dollar allocations. The effect on schools in the states was a 25% reduction in educational spending because of the decreased federal portion. The changes were outlined in the Education Consolidation and Improvement Act (ECIA), reflecting a shift in governance focus away from federal control (Natriello & McDill, 1999; Timar, 1994) through the reduced federal funding.

At the same time, state governments were given more flexibility to make decisions in several new policy areas (Fowler, 2000; Kaplan & O' Brien, 1991). Local school districts lost most of their control over education when states were given more discretionary power over federal education funds (Fowler, 2000). For example, local districts had previously controlled curriculum policy; however, during the 1980s states began to implement basic skills curricula and statewide testing programs. Even states with limited curriculum involvement intensified their control over the curriculum (Fowler, 2000; Porter, Archbald, & Tyree, 1991). During the same period of time, T.H. Bell, the Secretary of Education under President Reagan, created the National Commission on Excellence in Education (1983). The commission gave a report on educational quality in America. Based on the report of the Commission, *A Nation at Risk*, the president alerted the country that the schools were in trouble with the future and security of the country in jeopardy.

Leaders in every state began developing plans for school renewal including improving existing programs. Few questioned the deep structure of existing schooling, including uniformity of classrooms, existence of curriculum tracks, grouping of students by grade levels, standardization of curricula, reliance on textbooks, orientation of control and group instruction, and use of norm-referenced tests to measure success (Tye, 1992; Wallace & Graves, 1995). However, several scholarly papers supported school renewal through standards and assessment. One paper, which came from the Commission on the Skills of the American Workforce (1990), described standards and the use of standards for high school certificates. The standards based certificates would provide the basis for hiring workers and allowing students into post-secondary institutions (Rothman, Slattery,

& Vranek, 2002). In addition, the report became the basis for the New Standards Project—a consortium of states committed to developing K-12 standards and assessment (Rothman, et al., 2002).

Yet, standards based systems overlooked providing students with the opportunity to learn. Effective policies and resource systems supporting instruction would have provided students with this opportunity to learn (National Council on Education Standards and Testing, 1992; Rothman, et al. 2002; Smith & O’Day, 1990). Stedman (1996) used the content of algebra to criticize any inclusion of opportunity to learn standards. Stedman found low mathematics scores for students whether or not the students were taught algebra before or after testing. Berliner and Biddle (1996) also found that students had similar achievement levels, even when those who had and had not taken advanced math courses were separated. A court case in Florida, *Debra P. v. Turlington* (1979, 1981), provided the standard by which a fair opportunity to learn was legally defined. In that case, plaintiffs challenged use of a minimum competency test for graduation requirements. The court accepted evidence that the test measured skills included in the curriculum by showing that most teachers considered the skills ones they should teach. Thus, on the opportunity to learn question, the court found that students had a fair opportunity to learn (*Debra P. v. Turlington*, 1981; Heubert & Hauser, 1999). However, many believe that the test as a valid measure of opportunity to learn should be based on actual instruction (Airasian & Madaus, 1983; Le & Klein, 2002), affecting the validity of high stakes tests.

Large scale testing had previously been a part of the federal support of education. Testing had been taking place through Title I, and was limited to assessment of low

achieving children in poor neighborhoods (Heubert & Hauser, 1999). Testing was subsequently criticized for creating a *Lake Wobegon* effect; that is, states and districts were reporting that all their students were scoring above the national norm (Koretz, 1988; Linn, 2000). Reasons for above average achievement across states included the use of old norms, repeated use of the same test, exclusion of some students from testing, and focusing instruction on the skills tested (Koretz, 1988; Linn, 2000; Linn, Graue, & Sanders, 1990; Shepard, 1990).

A Nation at Risk impacted educational direction. The report listed four recommendations of strengthening graduation requirements, providing rigorous and measurable standards, increasing time in schools, and improving teaching. The goals would have remained somewhat unnoticed had the states not responded to them (Fowler, 2000; Ravitch, 2000). In 1986, the National Governors Association, with leaders such as Lamar Alexander of Tennessee and Bill Clinton of Arkansas, lent support to the recommendations (Ravitch, 2000). In 1989, six national educational goals were established which encompassed earlier recommendations. The national goals had now received state coordination and buy-in. Such state government involvement in educational policy was unprecedented in American history (Bowman & Kearney, 1986; Fowler, 2000; Mazzoni, 1995; Nathan, 1993). Of special interest was Goal Three in which no student would leave grades 4, 8, and 12 without demonstrating competency in challenging subject matter. The goal was to prepare students for more learning, citizenship, and employment in a modern economy (United States Department of Education, 1995). Obviously, testing beyond the Title I program evaluation was needed.

By the mid-1980s, widespread support for student assessment existed. Thirty-three states mandated some form of minimum competency based testing (Heubert & Hauser, 1999; Office of Technology and Assessment, 1992). Most testing during the 1980s involved multiple-choice items, though, in some systems, direct assessment of writing was also included (Hamilton & Koretz, 2002). Also enacted during this time was the Hawkins-Stafford Amendments in 1988, which reflected a shift in the emphasis of Title I monitoring from fiscal and programmatic compliance to educational improvements (Jennings, 1991; Natriello & McDill, 1999). Views that the achievement gap between less advantaged and more advantaged students caused the shift in emphasis prevailed (Kober, 1992; Natriello & McDill, 1999). The strategy for implementation of the national educational goals was called America 2000. The strategy provided an outline for goal implementation through raising academic achievement for all students, setting a target graduation rate, preparing young children for school, increasing adult literacy, and reducing substance abuse and school violence (Ravitch, 2000). By 1992, over 48 states and 2,000 communities had committed to America 2000. By the mid-1990s, 18 states had test-based requirements for high school graduation (Bond & King, 1995; Heubert & Hauser, 1999).

The 1990s included changes in testing that some called the second wave of reform (Hamilton & Koretz, 2002). It was caused by teachers teaching to the test, with an inflation of test scores and no accompanying increase in student learning (Hamilton & Koretz, 2002). In other words, test score improvement came from students being more familiar with the test through test preparation and less from enriched instruction. Reformers called for using new assessments that would avoid inflation of test scores

without an increase in student achievement. As a result, state reformers created tests that would provide quality instruction even when teachers target instruction toward helping students do well on the tests (Hamilton & Koretz, 2002). Testing format changed to include essays, short-answer questions, portfolios, and hands-on performance assessments, rather than a reliance on multiple-choice items (Hamilton & Koretz, 2002). While most state tests continued to rely on multiple-choice tests, some revised their tests to include extended-response or essay writing and/or other similar items. The state tests of Vermont and Kentucky (Education Week, 2001; Hamilton & Koretz, 2002) went further and began requiring portfolios of student work.

During the Clinton presidency (1993-2001), the administration pushed for laws to support implementation of the strategies articulated in America 2000. America 2000 became Goals 2000 and significant reform initiatives became law in 1994, as a part of the *Improving America's Schools Act* (Natriello & McDill, 1999). Each state educational agency was charged with developing content standards and performance standards (United States Department of Education, 1998). Content standards would detail what students should know and required a consensus on education for American children. Once consensus was reached, content standards were tied to performance measures. The consensus set up an accountability system, which provided “sanctions, supports and rewards for performance” (United States Department of Education, 1998, p.2).

When Bush replaced Clinton in 2001, content standards and assessment remained a focus. During the Bush administration, assessment became stronger through the *No Child Left Behind* (NCLB) legislation. Under NCLB, states assess students in all grades between 3-8. Annual school reports carry summaries of the testing disaggregated by race,

gender, and other criteria (NCLB, 2001). In effect, NCLB places control of accountability with the federal government (Jones, Jones, & Hargrove, 2003; Kiely & Henry, 2001). Thus, state accountability practices, including testing for nearly every public school child, became standardized to a federal format under NCLB.

Louisiana Response

Louisiana leaders responded to federal reforms by taking advantage of the availability of funding being offered to encourage SBR. The federal Department of Education, in collaboration with other federal agencies such as the National Science Foundation, offered funding for development of academic standards. One specific content standard funding source for Louisiana came in 1993 and provided the prompt for other content areas to follow. The funding was a three-year grant for \$900,000. The broader reform effort was called the Louisiana Systemic Initiatives Program (LaSIP). The mission of LaSIP was achievement of standards based reform in both mathematics and science education (Finley, 1999; LaSIP, 1997). LaSIP was the connection between reform in Louisiana and the national reform agenda (Finley, 1999). While the LaSIP effort contained a broader scope of teaching methodology and teacher content knowledge, the state also used this particular grant money to develop the first set of Louisiana content standards. Once completed, the state focused on other core subject areas of English language arts (ELA), social studies, foreign languages, and the arts. Content standards, one part of the reform initiative, became tied to performance standards to determine how well students met the standards. Previously, Louisiana linked assessments to competencies by Act 750 of 1979 (Breckenridge & Goldstein, 1998; Finley, 1999). New

reform efforts followed the earlier method of linking assessments to competencies, but added high stakes consequences for learners.

In 1996, Louisiana governor, Mike Foster, endorsed “setting high academic standards, [and] developing appropriate assessment” (Finley, 1999, p. 4). Within four years, Louisiana became the first state to require fourth and eighth graders to earn a successful score on a standardized CRT for promotion to the next grade (Amrein & Berliner, 2002; Johnson, & Johnson, 2002; Robelen, 2000). Thus, testing in Louisiana became high stakes for fourth and eighth graders. The governor had not acted alone. The Louisiana legislature responded to reform efforts. In 1997, the Louisiana legislature passed an act creating the School and District Accountability Commission and assigning it to ensure measures of student performance were in place. Hence, students, schools, and districts became accountable for student performance.

The response of Louisiana to reform efforts also includes how well SBR has been implemented. The accountability program of the state comprises more than having standards and assessment. Alignment of standards to the assessment is also needed. Because of the high stakes consequences of assessment in Louisiana, the standards need to be clearly aligned. The American Federation of Teachers (AFT) rates the alignment efforts of states against a group of indicators. In the rating system, the standards of Louisiana in mathematics and science across all grade levels showed clear alignment. Standards in social studies specified clear and specific understanding in middle and high school; however, English lacked clarity across all grade levels. Overall, the AFT (2001) assessed testing in Louisiana as being aligned to the standards, but the state rating reflected low scores on deriving clear meaning from the wording of the standards. Here,

lack of clear and specific content standards in English across all grade levels contributed to the finding (AFT, 2001). For a state leading in consequences tied to high stakes testing, not providing clear and specific content standards in an area relied on heavily for promotion and graduation purposes can be problematic. However, in the spring of 2004, the Louisiana Department of Education (LDE) released grade-level expectations (GLEs). The GLEs provide more specificity to standards and benchmarks. The GLEs detail standards based content that should be mastered at each grade (LDE, 2004). The implementation of GLEs began in August 2004.

The overall instructional plan for the state, *Louisiana Education Achievement and Results Now for the 21st Century* (LEARN) document, includes many objectives. One of the objectives calls for “high academic standards and appropriate assessments” (LEARN, 1996, p. 3). In addition, the manner in which high academic standards are implemented is contained in state policy. The policy for Louisiana, *LEAP for the 21st Century High Stakes Testing Policy*, provides high stakes criteria for students. The discussion here is limited to eighth graders because of the focus of the current study. For eighth grade promotional purposes, “a student may not be promoted to grade 9 until he or she has scored at or above the *Approaching Basic* level on the English Language Arts and mathematics components of the grade 8 LEAP 21” (LDE, 2003, p.5). Table 3 outlines student placement decisions for students taking LEAP 21 and either passing or failing the test. Students must still meet coursework and attendance requirements established by district and state policy. Students go to grade 9 by passing subtests of ELA and mathematics or under a retention limit policy.

For students who do not pass the test, three options are available based on the number of sub-tests failed, the students' attendance at summer remediation, and the student's age. Option 1 students repeat grade 8. Option 2 or grade 8.5 students take high school courses, but also take a remedial class in the failed sub-test content. The grade 8.5 student is located on a high school campus in the transitional grade, 8.5. Option 3 students pursue a Graduate Equivalency Diploma (GED).

Table 3

LEAP 21 Policy Requirements and Placements for Taking the Grade 8 Test

Placement	High Stakes Testing Policy Requirements
Grade 9	<ol style="list-style-type: none"> 1. Student passes ELA and mathematics subtests of LEAP 21. 2. Student repeats grade 8, fails one subtest, and attends summer remediation. (Student exceeds the retention limit.)
Option 1/Grade 8	<ol style="list-style-type: none"> 1. Student fails one subtest and does not attend summer remediation. 2. Student fails two subtests with or without summer remediation.
Option 2/Grade 8.5	<ol style="list-style-type: none"> 1. Student fails one subtest and attends summer remediation.
Option 3/ Pre-GED Skills Option	<ol style="list-style-type: none"> 1. Student fails one or more subtests, and the student is at least 16 years old. 2. Student voluntarily agrees to enter the GED program.

In summary, federal involvement in education changed, resulting in standards based reform (SBR) and federal provisions to have all students demonstrate competence in

specified gatekeeper grades. Louisiana responded by taking the lead in holding students accountable through high stakes promotional consequences. Because Louisiana, a state with a large minority population (U.S. Census Bureau, 2003a), took the lead in using assessment for promotional purposes (Amrein & Berliner, 2002; Johnson, & Johnson, 2002; Robelen, 2000) and because prior misuses of testing have disadvantaged minorities (Chachkin, 1989; Garcia & Pearson, 1993; Gamoran, 1996; Garcia et al., 1989; Rebell, 1989), a discussion of testing takes place in the next section.

Review of Research on Testing: Roles and Attributes

Responses to high stakes tests by the general public and professional organizations are given in this section revealing the sociopolitical context of the tests. An examination of the impact of high stakes testing on teacher methods and student learning follows, showing a lack of uniformity in best approaches in teaching and learning. Finally, a description of appropriate test use provides a gauge to review the high stakes tests of Louisiana and suggests how a disparate impact on poor students takes place.

Support and Challenges

In the U.S., the public supports high stakes testing despite evidence that it leads to lower rates of promotion and graduation (Heubert & Hauser, 1999; Hochschild & Scott, 1998; Johnson & Immerwahr, 1994). The public supports use of high stakes testing to identify student and teacher weaknesses, decide who is promoted, and assess the rank of a school (Heubert & Hauser, 1999). Researchers use the annual Phi Delta Kappan/Gallup poll to measure opinions by the public on high stakes testing. The support has been consistent over time since 1978 (Heubert and Hauser, 1999; Hochschild & Scott, 1998).

Challenges to the use of high stakes tests also occur. The National Council of Teachers of English (2000) criticizes the tests for making important student decisions based on one test rather than on multiple measures for promotion, graduation, or employment. The organization refers to such use as unethical and unsound. Other organizations take similar positions. For example, the College and University Faculty Association of the National Council for Social Studies calls for the use of authentic assessment, utilizing portfolios of student work. This group also opposes all uses of high stakes standardized tests (California Coalition for Authentic Reform in Education, 2002).

The American Education Research Association (AERA) presents a position statement on use of high stakes testing in PreK-12 education based on the 1999 *Standards for Educational and Psychological Testing*. The statement represents a consensus with the American Psychological Association (APA) and National Council on Measurement in Education (NCME). AERA (2000) opposes the use of a single test for decisions “that affect individual students’ life chances or educational opportunities” (p.2). The statement, according to AERA, represents the most authoritative guidelines for appropriate test use and notes that such tests should be only one part of information on which educational opportunities are based. The National PTA took a similar position concerning use of a single test for high stakes decisions. “At no time should a single test be considered the sole determinant of a student’s academic or work future” (National PTA, 1996, p.1).

While AERA opposes the use of a single test for high stakes decisions, the organization does include guidelines for use of a test as a single measure. As a minimum assurance of fairness, the guidelines provide for multiple opportunities to pass. Louisiana graduation high stakes tests offer multiple opportunities to pass as students begin testing

in grade 10, offering many opportunities to test and pass before graduation. However, for grade 8 promotional purposes, a student is assessed beginning in the spring of the school year. Another opportunity presents itself at the end of the summer. Therefore, only two opportunities are presented to pass before retention in grade 8.

AERA also recommends meaningful remediation in states that mandate high stakes tests. In Louisiana, summer remediation fulfills this requirement. AERA further recommends a sufficient time lapse between testing episodes. Eighth graders in Louisiana retest after summer remediation sessions. However, the summer period may not be a sufficient time lapse in between testing sessions; this depends on the individual student and the quality of the remediation during summer school. In addition, professional organizations and researchers call for on-going assessment of high stakes testing for intended and unintended effects (AERA, 2000; Linn 2000). The examination of the disparate impact of high stakes tests in the current study makes it a part of that on-going assessment.

Part of the misuse of high stakes testing occurs when one test fills multiple purposes. McDonnell (2002) recognizes that many states use high stakes tests for purposes such as providing information about the status of the system, influencing classroom instruction in a particular direction, rewarding and sanctioning schools, motivating students to perform better, and making decisions about student promotion and graduation. Using one test for multiple purposes puts policymakers at odds with the professional standards of the testing and measurement community (McDonnell, 2002). While policymakers find that having one test serve multiple purposes is cost efficient, the testing and measurement community

realizes that utilizing one test for multiple purposes exceeds the capacity of the test (Heubert & Hauser, 1999).

Hauser, Pager, and Simmons (2000) suggest that differences in retention rates of Black and White students can largely be explained by social and economic factors. However, differences in test scores are generally larger than what would be expected from social and economic differences. The difference suggests that tying test scores to promotion purposes has a disparate impact on racial and ethnic minority students (Hauser et al. 2000). Thus, the present study finds strong support in the literature for examining unintended consequences, including a disparate effect on minority students. While professional groups and supporters differ, state personnel play a central role in determining what and who gets tested.

What and Who Gets Tested

State officials determine content standards and constructs for inclusion in the standards. Which standards and constructs state officials select does matter (Linn, 2000; Willingham & Cole, 1997). The content selected for high stakes consequences also matters (Linn, 2000). For example, in the 1994 test administration of the National Assessment of Educational Progress (NAEP), girls outperformed boys by 14% in reading, and boys outperformed girls by 3% in history and 10% in geography. This disparity implies inclusion of a particular content can advantage a particular group of students (Linn, 2000). The manner in which test construction takes place also matters. Girls tend to earn higher scores than boys on short answer and essay questions, but the reverse is true for multiple-choice questions (Bolger & Kellaghan, 1990; Mazzeo, Schmitt & Bleistein, 1993). Here, the discussion of gender helps illustrate how state

selection of standards and constructs can advantage a particular subgroup and disadvantage others. This discussion is intended to suggest that high stakes tests, by construction, may advantage or disadvantage groups of students.

Current reform efforts place emphasis on the inclusion of all students (Linn, 2000), with Meyer (1996) advocating data collection of family and community characteristics, to which test scores would be tied. Thus, the data collection would explain disadvantages of groups of students. Other researchers (Elmore, Abelman & Fuhrman, 1996; Linn, 2000) argue a dichotomy in score computation. The dichotomy occurs when school personnel become responsible for factors beyond their control such as socioeconomic status (SES). A problem also occurs when school personnel do not accept responsibility for factors within their control, acting to institutionalize low expectations for poor and minority students. Critics of Meyer find the practice of data collection of family and community characteristics burdensome. However, the data can be considered as a measure of increased validity.

Meier (2002) criticizes use of cutoff scores for doing what psychometricians said could not be done—presenting the score of a student as an absolute value of learning. Meier presents a grade equivalent score of 4.5 and establishes that the score does not mean that the student reads as a fourth grader in the fifth month but somewhere between 3.9 and 4.9 in most cases. However, conferring an absolute value to the score provides the basis for making graduation and promotion decisions. State personnel set cutoff scores in several ways. Horn, Ramos, Blumer, and Madaus (2000) identify three ways; Modified Angoff method, contrasting groups, and bookmark method. All of the methods involve subjectivity as all rely on predictions about how students perform on test

instruments. Heubert and Hauser (1999) referred to setting cutoff scores as “ inherently judgmental” (p.99), no matter which method test developers select.

In the Modified Angoff method, predictions consist of percentages of students successfully answering each test question. In contrasting groups, student grouping occurs based on predictions as to whether students score well or less well. Actual performance and predicted performance comparisons take place. In the bookmark method, predictions take place at the point at which more capable students correctly answer test items than less capable students. Advantages of Modified Angoff and bookmark methods include setting cutoff scores without the content experts knowing the students. The advantage assists state committees of content experts. The contrasting groups method requires the experts to know the students, thereby reducing its usefulness with a committee of state content experts. The bookmark method allows several points to be identified on a continuum with levels of proficiency set.

Louisiana officials selected curriculum specialists and teachers as content judges. These officials and judges selected the bookmark method to set cutoff scores for eighth graders on LEAP 21 (LDE, 2001b). They also identified the achievement levels of *Advanced, Mastery, Basic, Approaching Basic, and Unsatisfactory* consistent with those used for the National Assessment of Educational Progress (NAEP).

According to the bookmark method, all assessment items were ranked from easiest to most difficult. Three groups of approximately 10 judges then bookmarked levels of proficiency, identifying the points at which students at each achievement level would correctly answer the test items. For example, judges asked themselves, “If a student were performing at the Basic level, would the student answer this item correctly”? Judges were

provided opportunities to reconsider their decisions and facilitators led discussions to clarify the achievement level selections. For Louisiana eighth graders, the politically acceptable passage rates for LEAP 21 became those established by NAEP test administrations. In effect, educators sought to match LEAP 21 achievement levels to those applied in NAEP scoring. Of interest, NAEP test administration allows exclusions for Limited English Proficient (LEP) students and students with an Individualized Education Plan (IEP) (Amrein & Berliner, 2002). However, all Louisiana eighth graders are held to the established cut scores on LEAP 21, except for 1% of eighth graders who are excluded because of severely disabling conditions.

Jaeger (1989) looked at competence being a continuous variable; however, selecting cutoff scores separates students into competent and incompetent (Heubert & Hauser, 1999). The separation of students takes place in Louisiana as students pass the test by receiving an Approaching Basic score or above or fail the test by receiving an Unsatisfactory score. In effect, the separation is passing or failing the test for the student.

Schools also receive labels based on student performance measures. Linn (2000) presents several scenarios as to what options states have in choice of data source and summary statistic. The discussion is included here to suggest that arbitrariness in selecting cutoff scores for students also takes place in assigning labels to schools. A study by Clotfelter and Ladd (1996) illustrated that data source and summary statistic selection matter by using a number of correlations in which the data source and summary statistic were varied. In the study, summary statistics such as school mean, mean gain, or residual score with adjustments for SES were examined for the level of correlation. The correlations ranged from .22 to .94. The range in correlations revealed the arbitrariness of

selection by state officials of criteria for assigning school labels in district accountability. Thus, a school labeled as deficient in one system could be labeled acceptable in another depending on the selected data source and summary statistic.

Testing Impacts on Classroom Practices

Beyond critics and advocates, researchers look at the impact of tests on teacher practices and methodologies and student achievement. Overall, research supports use of tests for improving classroom practices (Corbett & Wilson, 1991; Herman & Golan, 1993; Heubert & Hauser, 1999). However, improved classroom practices as a response to high stakes testing finds little support in improving student learning (Heubert & Hauser, 1999; Mehrens, 1998). The lack of support for improved student learning takes place even with teachers reporting they are working harder and focusing more on instruction (Finnigan & Gross, 2001; Kelley, Odden, Milanowski & Heneman, 2000). In addition, a small number of researchers report (Firestone, Mayrowetz & Fairman, 1998; Grant 2001) little to no effect on changes in teaching methodologies.

As to which particular teacher methodologies benefit student performance on high stakes tests, the research covers a wide range of methodologies and offers no clear benefit of any one method over the others. Teachers report using more worksheets, exercises, and drills (Calkins, Montgomery, & Santman, 1998; Jones, et al., 2003). In addition, researchers documented teachers adjusting their style and content to more closely match that which they perceive results in higher test scores (Dorn, 1998; Madaus, 1988, 1991; Shepard, 1991; Smith, 1991; Smith & Rottenberg, 1991). In a survey of 2,200 mathematics and science teachers, and with visits to six urban sites, researchers found that teachers in low-income settings are most likely to teach to the test (Center for the

Study of Testing, Evaluation, and Educational Policy, 1992; Garcia, & Pearson, 1993; Rothman, 1992). Mixed results came from North Carolina where teachers reported either increasing or decreasing their use of projects, lecture, textbooks, and worksheets (Jones, Jones, & Hargrove, 2003; Jones, Jones, Hardin, Chapman, Yarbrough, & Davis, 1999). Again, teachers selected their instructional methods on how they perceived the method would increase student test scores.

Jones, et al. (2003) suggest the reason for conflicting results in the research is that much of it is based on interviews and self-reports. They present four factors to explain why research findings are mixed. Factors influencing the type and amount of change teachers make in their methodologies include the type of high stakes assessment, the type of professional development offerings, the subject area tested, and the level of achievement at the school.

Beyond methodologies, teachers involved in preparing students for high stakes testing feel more pressure and responsibility (Jones, et al. 2003), work more effectively (Borko & Elliot, 1999; Wolf & McIver, 1999) and use assessment as an opportunity to rededicate themselves (Wolf, Borko, McIver & Elliott, 1999). As to the curriculum, research supports the curriculum becoming tighter, comprehensive, and consistent with standards (Borko & Elliot, 1999; Jones, et al., 2003), leading teachers to report a negative impact on their creativity (Gordon & Reese, 1997; Jones & Johnston, 2002; Wideen, O'Shea, Pye, & Ivany, 1997).

Negative impacts of high stakes tests can be justified if the tests lead to increased levels of student learning. Amrein and Berliner (2002) examined the impact of high stakes testing by examining achievement results using other indicators. The researchers

used the National Assessment of Educational Progress (NAEP), the American College Test (ACT), the Scholastic Achievement Test (SAT), and the Advanced Placement (AP) tests to assess high stakes testing impacts across 18 states with severe consequences attached to testing. In all but one analysis, student achievement was indeterminate, remained the same, or went down. The study included Louisiana. In charting student achievement trends in Louisiana, Amrein and Berliner examined NAEP data and found Louisiana students (a) gained 6 points on the nation in grade 4 mathematics from 1992 to 2000, (b) gained 0 points in grade 4 reading from 1992 to 1998 thus following the national trend, and (c) gained 13 points in grade 8 mathematics from 1990 to 2000. The results indicate the need for current data that tracks learner achievement through inclusion of promotional high stakes tests. A question arises concerning whether growth can be attributed to the inclusion of high stakes tests or whether scores were already improving.

One of the weaknesses of using NAEP data is that the data pertains only to grades 4 and 8. Louisiana students had high stakes consequences for graduation beginning in 1989 and for promotion in grades 4 and 8 in 2000. NAEP data needs to be gathered beyond 2000 for assessment of high stakes impacts for promotional purposes. Using grade 4 and grade 8 data to assess the impact of a graduation exam presents limited usefulness. Another weakness of using NAEP data includes student exclusions. NAEP staff randomly selects school districts and schools in states that choose to participate. In addition, student selection within the schools occurs randomly; however, students with Individualized Education Plans (IEPs) and students classified as Limited English Proficient (LEP) receive exclusions from testing. Amrein and Berliner (2002) find that

the exclusions bias NAEP results since Louisiana excluded 6% of special education students in 1994 compared to 13% excluded in 1998 (Education Reporter, 1999). Thus, improvements in NAEP potentially reflect percentage of exclusions rather than improvement of student learning (Amrein & Berliner, 2002).

One NAEP data collection procedure, which presents more valid information than others, avoids having to account for percentages of exclusions from year-to-year. The procedure tracks students over time. Scores from cohorts of students from grades 4-8 during 1994 to 1998 revealed Louisiana gained 6 points on the national average in reading, and for the period 1996-2000 the state lost 2 points on the national average. However, the promotional high stakes testing policy of Louisiana became effective in 2000 making the data less useful in telling whether high stakes tests caused NAEP scores to increase. Nevertheless, the data were useful in giving the improvement trend of Louisiana without high stakes testing; Louisiana was making progress, especially in reading.

Amrein and Berliner (2002) used other measures of achievement to chart the progress of Louisiana students. The researchers found, Louisiana students (a) lost .2 point to the nation on ACT from 1990 to 2001, (b) gained 19 points on SAT from 1990 to 2001, and (c) gained 2.6 percentage points in AP exams from 1995-2000. Each measure of achievement can be somewhat problematic. For instance, participation rates increasing or decreasing in a state reflects motivation to attend college, rather than actual achievement gains. Each of these measures is affected by such motivation (Amrein & Berliner, 2002).

In all, achievement trends across different measures indicate mixed progress for Louisiana students. The strongest evidence for achievement growth taking place in the

state came from the cohort trend in reading. Data representing changes in exclusions from yearly testing and student motivation allowed less clear indications of the progress of Louisiana students. At best, much of the data presented by the indicators represent a baseline against which to measure the influence of high stakes testing for promotional purposes.

Testing Impacts on Students

Students are affected by high stakes tests beyond the test-taking event. Tests negatively impact students through grade retention, student self-esteem, overage status, physical responses, and quality instruction. Table 4 provides a summary of the research on the negative impact of high stakes tests on students.

In looking at retention from 1971 through 1996, The Committee on Appropriate Test Use used birth cohorts to see the evolution of age-grade retardation or enrollment below the modal grade level for the age of a child. Of importance is that the differential between groups is barely noticeable at school entry; however, by age nine there is a hierarchy established. The hierarchy favors Whites and girls relative to Blacks or Hispanics and boys. By ages 15 to 17, “close to 50 percent of Black males have fallen behind in school” (Heubert & Hauser, 1999, p.122).

Franklin, Pernici, and Yuan (2001) examined student retention rates in the state of Louisiana during the period 1997-98 to 2000-01. The K-12 retention rate of Louisiana, holding around 8%, rose during the period 1997-2000 by less than 1%; however, the rate climbed to 11% with high stakes testing. Also, the study revealed that Black students were retained at a higher rate than other ethnic groups. Amrein and Berliner

Table 4

Negative High Stakes Test Impacts on Students

Effect	Evidence
Grade	Retention mostly impacts Black males (Heubert & Hauser, 1999).
Retention	Louisiana Black students are retained at higher rates than other groups of students (Franklin, Pernici, & Yuan, 2001). Higher rates of retention take place in Louisiana with high stakes testing (Amrein & Berliner, 2002).
Student Esteem	Negative self-esteem results from retention (Gampert, 1987; Gottfredson, Fink & Graham, 1994; Heubert & Hauser, 1999, Roderick, Jacob, & Bryk, 2002).
Overage Status	Increased dropout rates occur as students become overage, especially for low socioeconomic students (Gampert, 1987; Gottfredson, Fink & Graham, 1994; Heubert & Hauser, 1999; Reardon, 1996; Roderick, Jacob, & Bryk, 2002).
Physical Responses	Testing causes headaches, irritability, aggression, freezing-up, and vomiting (Hoffman, Asaf & Paris, 2001); vomiting and crying observed during Louisiana test week (Johnson & Johnson, 2002).
Quality Instruction	Low performing schools have the least qualified teachers (Danielson, 1999; Jones et al., 2003; Wayne, 2002); low-test performance means minority students lack quality instruction rather than test score bias (Popham, 2000). Black students receive narrow focus of test preparation instruction (Jones et al., 2003; Kohn, 2000; McNeil, 2000).

(2002) document 15% of the Louisiana fourth and eighth graders repeating in the fall of 2000 and 10-15% being retained in the fall of 2001.

Critics of retention policies argue that retention negatively impacts dropout rates and increases negative academic self-esteem (Gampert, 1987; Gottfredson, Fink & Graham, 1994; Heubert & Hauser, 1999; Roderick, Jacob, & Bryk, 2002). Also, Reardon (1996) used data from the National Educational Longitudinal Study (NELS) to suggest that the presence of high stakes, grade 8 tests is associated with higher dropout rates, especially for schools serving mainly low-SES students (Heubert & Hauser, 1999).

Taking high stakes tests has also been shown to affect students physically. Hoffman, Assaf, and Paris (2001) surveyed 200 Texas teachers in examining the effects of the Texas Assessment of Academic Skills (TAAS) on students. Students physically respond to the tests with headaches, stomach upsets and vomiting, irritability, aggression, and freezing-up. Similar physical effects on students, vomiting and crying, were found in Louisiana schools during test week (Johnson and Johnson, 2002).

Research further supports concerns of principals about their ability to recruit teachers (Danielson, 1999; Jones, et al. 2003) in lower-performing schools, with an impact of the most needy students being taught by the least qualified teachers (Jones, et al., 2003; Wayne, 2002). Following in the thinking of quality instruction as being important, Popham (2000), in examining minimum competency exams of the seventies and eighties, suggests that students lack needed instruction when high percentages of minority students score lower on test items. Thus, as applied to SBR, quality instruction by qualified teachers is needed for minority students in preparing for high stakes tests; however, instruction for minority students focuses on test preparation to a greater extent than White

students (Rothman, 1996; Soloman, 1998). Kohn (2000) described the instruction received by minority students as practice in marking bubble sheets rather than pursuing “engaging projects that promote sophisticated learning” (p. 47).

Not all research supports negative test impacts on students. Research supports school based accountability programs including testing resulting in higher achievement for students (Ladd, 1999; Roderick, et al., 2002). In looking at the impact of high stakes tests on student achievement in gatekeeper grades, Roderick et al. (2002) examined school records of students in Chicago. The researchers found students with the lowest academic skills showing the largest gains on reading test performance in grades 3 and 6. In grade 8 all groups of students showed positive testing effects for reading; however, the opposite was true with mathematics. Higher achieving students showed the greatest gains in mathematics test performance in grades 6 and 8. For grade 3, high-risk students showed the greatest gains under the high stakes testing policy. The Chicago study used the Iowa Test of Basic Skills, a test developed in Iowa with a diagnostic purpose. Once again, the selection of data source is an important decision for a state or district, as illustrated by Clotfelter and Ladd (1996).

A discussion of the impact of testing on students that increased rates of retention took place in this section. An additional consideration involved the ability of principals to recruit teachers to provide quality teaching and build capacity in poorer districts. Finally, a discussion of how high stakes testing impacts student learning in Chicago gatekeeper grades documented how school reform impacts students. Discussion more focused on the Louisiana test takes place in the next section.

Appropriate Test Use

The National Research Council lists three criteria of validity, attribution of cause, and effectiveness of treatment for judging appropriate test use. The criteria provide a guide for high stakes test use (Heubert & Hauser, 1999). The three criteria, when applied to LEAP 21, require individual analysis.

Assessing whether the high stakes test is valid requires answering that the test is appropriate for the particular purpose (Heubert & Hauser, 1999). The purpose of the reform initiative in the state of Louisiana was to end social promotion, “the practice of passing students to the next grade even if they do not have the skills needed to succeed” (LDE, 2000a, p.3). At a minimum, the high stakes test of Louisiana ends social promotion for students not meeting the cutoff score; however, determining whether failing students have the skills needed to succeed requires multiple measures of their skills (AERA, 2000).

For attribution of cause, the performance or skill of the student must be based on whether the student received appropriate instruction (Heubert & Hauser, 1999). Appropriate instruction is being addressed through reform efforts in improving teacher capacity by defining a *highly qualified* Louisiana teacher. The reform effort mandate came from the federal NCLB (2001) legislation. While Louisiana is one state in which state officials provide information to parents concerning the credentials of every public school teacher in the state, the state received only a C+ in efforts to improve teacher quality (Quality Counts, 2003). The grading involved measures of teacher assessment, teaching in the field, professional support and training, and teacher education. Louisiana

falls short under attribution of cause by holding students accountable for test performance while providing only C+, or adequate efforts in providing quality teachers.

The last criterion for judging appropriate test use requires answering whether test performance leads to educationally beneficial decisions (Heubert & Hauser, 1999). Answers to this question come from the test itself. Retention has been shown to have negative impacts (Gampert, 1987; Gottfredson, Fink & Graham, 1994; Heubert & Hauser, 1999; Roderick et al., 2002), and data from NAEP (Amrein and Berliner, 2002) provide some evidence of improved achievement. However, more data needs to be gathered to thoroughly answer this question. Through the first phase of the current study, a description of the test score differences for Black and White students were given, and in the second phase, parental perspectives were used to explore whether test performance has led to educationally beneficial decisions for their children.

Heubert and Hauser (1999) list criteria in assessing whether test use in high stakes decision-making is illegally discriminatory. Using the criteria in assessing discriminatory effects would necessitate determining whether test use has no intent to discriminate, no effect of prior discrimination, no impact of being educationally unjustifiable, and no alternatives with less disproportionate impacts. The first criterion raises doubt for reasons presented earlier concerning stronger accountability systems in states with high percentages of minority populations (Carnoy & Loeb, 2002). States with high percentages of minority students use high stakes tests even though minority students do not test as well as White students on these tests (Gladfelter, 2000; Lewis, 2000). As to the second criterion, the high stakes test implementation in Louisiana raised concerns of prior discrimination by allowing minority students to be held to the same standards without

adequate instructional materials and climate for learning (Johnson & Johnson, 2002). More time is needed to assess whether the impact of high stakes testing is educationally justifiable. The discussion of NAEP scores and other achievement indicators previously presented revealed that current data is needed to assess positive impacts on these measures of student achievement due to the use of high stakes tests. These measures of achievement should improve if use of high stakes tests is educationally justifiable. Finally, authentic assessment has been suggested as an alternate with less disproportionate impacts on groups of students in some states (Washor & Mojkowski, 2003). However, when more subjective measures such as authentic assessment have been used, concerns arise about bias and non-objectivity, which lead to formalizing and standardizing (Madaus & Tan, 1993; Resnick, 1981).

The previous section provided discussion of stakeholders and their positions as advocates or proponents of high stakes testing. In addition, discussion took place showing the manner that state officials select standards and constructs matters. How state officials set cutoff scores for students, more specifically how Louisiana officials set cutoff scores, was provided in this section. Also, research was presented showing arbitrariness is not limited in holding students accountable, but the arbitrariness also applies to accountability labels a school receives. The impact of testing on classroom practices and the manner in which instruction is delivered was discussed in this section. The discussions were followed by research showing trends in learning through indicators such as NAEP. Next, impacts of high stakes testing on students were provided. The impacts were described beyond a test score revealing how students are affected by the test. Lastly,

a discussion specific to the test in Louisiana took place. The next section provides a discussion of the lens to view SBR implementation.

Related Research on Chaos Theory

The current section begins with descriptions of theories that researchers use to describe SBR implementation. The first two, the Heisenberg Uncertainty Principle and complexity theory, present features that provide an explanation for parts of SBR; however, the use of each theory lacks comprehensiveness. Chaos theory provides a lens for explaining rapid changes in school governance in recent years (Griffiths, Hart, & Blair, 1991; Maxcy, 1995), ultimately impacting classroom teaching and testing consequences for students. Chaos theory is used to describe SBR within and across different sociopolitical forces; however, discussion begins with other possible frameworks for understanding SBR.

Heisenberg Uncertainty Principle

In examining high stakes tests across 18 states, Armrein and Berliner (2002) use the Heisenberg Uncertainty Principle to explain attaching personal and educational consequences to test performance. The principle proposes that the more important any quantitative measure becomes in social decision-making, the more that measure will corrupt or distort the social behavior intended for monitoring. Thus distortion and corruption occur to inferences based on quantifying social behavior. For example, in looking at the score on a standardized test for a student, one distortion would be that the score represents training in test preparation rather than learning through an enriched curriculum. The distortion comes from narrowing of the curriculum (Madaus, 1988, 1991; Shepard, 1991; Smith, 1991; Smith & Rottenberg, 1991) by using item teaching in

test preparation. Thus, use of the Heisenberg Uncertainty Principle may be useful in understanding the disparate effects of high stakes tests through use of distortions.

However as will be shown later, SBR involves more than distortions.

Complexity Theory

O'Day (2002) utilizes complexity theory to explain interactions between organizational layers involved in school accountability. Complexity theory, more closely aligned to chaos theory than the Heisenberg Uncertainty Principle, acknowledges a non-linear structure in which accountability (a) influences from outside what happens inside schools, (b) recognizes the school as the unit of intervention and the individual as the unit of action, and (c) holds the importance of information as essential to school improvement (O, Day, 2002). Complexity theory as applied to accountability recognizes many interactions. Some interactions include teacher-student, administrator-teacher, and superintendent-administrator; the outcome of each interaction is difficult to predict. The interactions reflect interdependence and change, as well as different interpretations to each participant. Even the manner in which district personnel convey expectations to school level personnel presents a variation in reform implementation across districts (Spillane, 2000, 2002; O'Day, 2002). The variation presents further evidence of the complexity involved in interactions ultimately impacting students held to high stakes consequences. O'Day (2002) acknowledges that teachers lack a response of working toward school improvement because of the top-down, well established, organizational system of school bureaucracies. The bureaucracies also place low-performing schools at a disadvantage by threatening them with outcome based accountability sanctions, thus adding to the stress of educators and students. O'Day (2002) suggests a combination of

professional and bureaucratic accountability that would be useful for fostering school improvement. Adding the dimension of professional accountability for teachers would positively impact the capacity of teachers to provide meaningful instruction. Teachers providing meaningful instruction would aid students held to high stakes consequences. While complexity theory is useful in understanding the organizational layers impacting high stakes testing, chaos theory presents a broader lens for the implementation of high stakes testing.

Chaos Theory

The substance of SBR implementation examined through a lens of chaos theory allows understanding of the school operation as multi-leveled, complex, and disorderly. In addition, chaos theory allowed recognition of SBR as a part of a recursive and dissipative cycle, which was presented in Chapter One.

Chaos theory explains SBR as multi-leveled. For instance, school districts respond to state and federal policies and influence the implementation of these policies (Goertz, Floden & O'Day, 1995; Jennings & Spillane, 1996; Spillane, 1998). Although different ideological beliefs exist between federal and state policies, districts implement these policies. Because these policies tend to be ambiguous, school districts implement them in ways that advance local agendas (Firestone, 1989; Spillane, 1998). Examination of SBR through use of chaos theory allows understanding of the ways that different layers of educational organization in a school system reinterpret reform policies and thereby add to the chaos-order status.

In addition, schools are organizations and as such have multi-layers of operation through services to students and teachers. The bureaucracy of the school organization has

been described by some as loosely coupled (Bidwell, 1965; Maxcy, 1995; Weick, 1976); that is, each layer works with little consideration or direct influence by others. Some researcher's view loose coupling as necessary in an organization full of unique, daily, human interactions; however, a loosely-coupled school organization adds to the challenges of implementing SBR. In SBR, the student is held accountable in a system that functions with many units lacking focused direction (Shipps & Firestone, 2003). Swanson and Stevenson (2002) refer to loose coupling in a system attempting to enact change as "dysfunctional" (p. 2). The alternative becomes for the school organization to reflect a more tightly coupled system. In such a system, the school as an organization becomes too machinelike by de-emphasizing human interactions (Maxcy, 1995; Simon, 1947, 1965).

In SBR, the range of observers expands making SBR implementation complex. Complexity is also a descriptor of chaos theory. Information about test scores and failure rates previously shared among state, district, and school personnel becomes published and disseminated to parents and community members through school report cards, newspapers, and websites (Dorn, 1998). The complexity occurs through transitions as misinformation expands, with understanding of the observers limited by their experiences (McDonnell, 2002). For instance, parents want their child to learn; however, the measure of learning for the child becomes passing the test. Misunderstanding surrounds the test as an accurate measure of student learning depending upon the knowledge and experiences of the observers. Ultimately, however, the only indicator of learning for the parent becomes whether or not the student meets the cutoff score and passes the test.

School performance, as measured through annual testing, becomes the political agenda of politicians and the perception of community members of instruction taking place at particular schools (McDonnell, 2002). Moreover, standards and assessment for some politicians become preserving public education in the face of moves to implement a voucher system (McDonnell, 2002). In effect, holding students to high stakes consequences serves as an alternative to vouchers. Politicians rely on scores to convey reality to the public when, in effect, scores are “like a thermometer, but really tell[ing] nothing about the health of a school” (Goodlad, 1984, p.14). The organization becomes more than a sum of parts. In addition, chaos theory recognizes the organization as complex rather than simple. Most importantly, chaos theory recognizes the organization as constantly fluctuating. In effect, chaos becomes symptomatic of a complex system (Gleick, 1987; Maxcy, 1995) as order is transitional. Maxcy (1995) uses a metaphor to describe the reality of a complex system as “a pulsating chaos-order machine, . . . a given and predetermined state of reality” (p.38). Thus chaos is a predetermined state of reality with fluctuations between more chaos and less chaos; however, chaos is always present.

Disorderliness, a third descriptor of chaos theory, applies to the substance of SBR development. Reform takes place in having all students meet high academic standards (Linn, 2000; No Child Left Behind, 2001). Also, reform occurs in correcting prior abuses of testing-namely, disproportionately sending more Black students into special education (Chachkin, 1989; Garcia & Pearson, 1993) and lower tracks (Gamoran, 1996; Garcia et al., 1989; Rebell, 1989). However, SBR, through high stakes testing, impacts Black students to a greater extent because Black students do not perform as well as White students on the tests (Gladfelter, 2000; Lewis, 2000). In effect, while reform efforts

should ideally be advantageous to Black students, life chances for Black students are actually more negatively impacted by retention policies of high stakes testing when the test is used as a measure for school accountability. Thus, disorderliness occurs when the effect of the reform is not met for Black students, as measured through a disproportionate failing of the test.

Chaos, or disorderliness, is further shown by the shift in curriculum away from teachers to politicians, leaving teachers caught between best practices and a need to show what their students can do on the test (Jones et al., 2003). Teaching to the test becomes common practice for teachers, with teachers of minority students engaging in test preparation practices at higher rates than teachers of non-minority students (Jones et al., 2003; Kohn, 2000; McNeil, 2000). Students, who are identified as at-risk of failing a high stakes test or who have failed the test, attend remediation classes; also many teachers become involved in item teaching test preparation (Madaus, 1988, 1991; Shepard, 1991; Smith, 1991; Smith & Rottenberg, 1991). Goodlad (1984) finds the practice detrimental because it denies slower students access to learning activities in other fields enjoyed by able students. Furthermore, disorderliness is suggested by the state holding schools accountable for test scores and being granted constitutional authority to take over failing schools. Goodlad (1984) would have states hold districts accountable rather than individual schools. Goodlad (1984) refers to the state holding principals, teachers, and individuals accountable as “unrealistic and ultimately damaging” (p.274). Accountability impacts schools; however, schools, bound by district and state policy, possess little autonomy (O’ Day, 2002; O’Day & Smith, 1993).

Much of the present chapter focused on the shift of educational control between the levels of government (Fowler, 2000) and its impact. The shift directly impacted teaching and ultimately impacted how achievement was defined, namely, by meeting a cutoff score. Cutoff scores for high stakes tests were not set by test-makers. Political officials set the scores in state departments of education, making the scores political rather than technically normed (Meier, 2002). However, choosing a cutoff score should involve social and political concerns (Ellwein & Glass, 1989; Heubert & Hauser; 1999). McDonnell (2002) recognizes standards and assessment to be political issues, requiring judgments about values and touching upon philosophical and cultural debates. The debate involves “who should be held accountable to whom for what” (McDonnell, 2002, p.104). While not being presented in the chapter, some would suggest disorderliness of the aims of SBR measured through testing is inconsistent with other relevant educational issues. For example, SBR is inconsistent with socialization for students in a changing world (Darling-Hammond, 1992; Dorn 1998).

Beyond a description of SBR as multi-leveled, complex, and disorderly, the lens of chaos theory allows suggestion of a critical pragmatic view of SBR. To avoid SBR implementation becoming too machinelike or too loosely coupled, Maxcy (1995) suggests a critical pragmatic view away from traditional, bureaucratic, and statistical conceptions of public education. A critical pragmatic view allows local differences in schools to be “encouraged rather than expunged” (1995, p.55). The view engages a consistent pragmatic approach to schools, reducing the chaos-order pulsating effect of SBR implementation.

Overall, chaos theory represents the broader lens for studying the disparate impact of high stakes tests. Chaos theory allowed understanding of the different levels of the organization of education. Loose coupling and movement toward tighter control in SBR paralleled the pulsating chaos-order description of chaos theory. Expansion in the range of observers expanded the complexity of the system and the impact of SBR was shown in several ways: (a) appearing to disparately impact Black students when Black students should benefit from reforms, (b) shifting control of the curriculum from teachers to politicians, (c) holding schools accountable rather than districts, and (d) recognizing standards and assessments to be political issues. Therefore, chaos theory provided a lens needed for the current study.

Summary

Chapter Two provided a review of literature needed in a study of the disparate impact of high stakes tests. Federal educational financing changes beginning with the New Federalism of the Reagan administration allowed more state flexibility in education decision making, which resulted in less local control. State officials began making school renewal initiatives, including setting high academic standards, which were necessary for national security. From the 1980s through the 1990s, support grew for assessment of state standards through testing. Into the 2000s, standards and assessments received strong federal support. Louisiana officials developed standards using federally supported initiatives and in 1996 became the first state to tie promotion to a test.

The state continues the effort to make standards clear and specific and for students taking the test, several placements are possible including a transitional grade, 8.5. The public has supported the use of tests consistently over time. Many professional

organizations oppose the use of a test as a single criterion when the test affects a student's life chances and when the test is used for multiple purposes. The multiple purposes exceed the capacity of the testing instrument.

State officials select the content, which is tested, and the types of questions that are used on the test. The selections impact groups of students differently with the reform including all students meeting the same standards. Louisiana uses the benchmark method to set cutoff scores.

Research supports testing as improving classroom practices. Teachers select methodologies they perceive will raise test scores. Improvements in student achievement because of the promotional test remain questionable because use of academic gauges such as NAEP can be problematic. Students are negatively impacted by high stakes tests through grade retention, student esteem, overage status, physical responses, and quality instruction. Research also supports higher student achievement because of testing. State officials are making progress to ensure that LEAP 21 use meets the criteria for judging appropriate test use and that use of the test is not discriminatory.

Chaos theory is described as a lens to view SBR implementation. The lens allows an examination of the multi-leveled organization of reform. Multiple observers make the implementation complex and disorderly and the curriculum shifts away from classroom teachers toward the state holding schools accountable for state standards through student test scores. A critical pragmatic approach was suggested as a means to avoid machinelike school organizations when SBR becomes too orderly and to avoid dysfunctional ones when implementation becomes too loose. The next chapter details the methodology of the current study.

CHAPTER 3. METHODOLOGY

The present study examined forces and events involving high stakes testing for promotion in Louisiana. Scores of eighth graders on the criterion referenced test (CRT), *LEAP For the 21st Century* (LEAP 21), were examined to determine whether a disparity exists in the achievement of Black and White students, though evidence suggests Louisiana is similar to the rest of the country in the effect testing places on Black students as opposed to White students (Gladfelter, 2000; Lewis, 2000). A second phenomenon studied was perspectives of parents of students who took the test.

The study was conducted in two phases, consistent with the two phenomena. The design of each phase and the relevant research questions are discussed in the present chapter. For a complete listing of the questions that guided the study, see Chapter One, page 18. Answering the questions guided a mixed methodology (Creswell, 1998) or *quant + QUAL*, dominant-less dominant sequential study design (Tashakkori & Teddlie, 1998).

Phase I

Phase I explored whether differences exist in the test performance of Black and White eighth graders who lived in one of three community types; rural, suburban, or urban. The tests of interest are the English language arts (ELA) and mathematics subtests of LEAP 21. Students receive a scaled score for performance on each subtest while the state sets a minimum, or cutoff score, for the subtests determining whether or not a student passes to the ninth grade. Also, students are placed into one of five achievement levels for each subtest; *Advanced, Mastery, Basic, Approaching Basic, and Unsatisfactory*. Thus, test performance comparisons between Black and White students were made in three ways;

scaled score differences, pass/fail differences, and achievement level differences. The research questions for this phase were:

1. What are the differences in LEAP 21 ELA and mathematics scaled scores between Black and White eighth graders in rural, suburban, and urban community types?
2. What are the differences in LEAP 21 ELA and mathematics pass/fail rates between Black and White eighth graders in rural, suburban, and urban community types?
3. What are the differences in LEAP 21 ELA and mathematics achievement levels between Black and White eighth graders in rural, suburban, and urban community types?
4. What are the differences in LEAP 21 ELA and mathematics passage rates between Black and White eighth graders in rural, suburban, and urban community types when magnet school students are included versus when they are not?

Statistical analyses used in Phase I involved two levels of race, Black and White, and three levels of community type; rural, suburban, and urban. Community type was established using locale codes of the common core of data, part of the primary database on public elementary and secondary education in the United States (NCES, nd).

Independent variables of race and community type were used to analyze the dependent variables of grade 8 LEAP 21 subtest scores in ELA and mathematics.

Scaled scores for students testing in the spring of the 2002-03 school year in both ELA and mathematics came from the Division of Student Standards and Assessments of the Louisiana Department of Education (LDE). The LDE previously used these data to report student test results and for school accountability purposes. Thus the confidence in use of the data in the current study is increased because of the LDE use. The data were obtained by submitting a proposal to the LDE for the Louisiana 2003 student file. Once obtained, the data were organized utilizing a spreadsheet and imported into SPSS (SPSS, 1998) software for analyses.

A multivariate analysis of variance (MANOVA) of the scaled scores was used to find out whether differences by race of the students in performance on ELA and mathematics subtests were statistically significant across community types. Two assumptions for conducting MANOVA were met including independence of observations (Newton & Rudestam, 1999) and multivariate normality (Hand & Taylor, 1987). The third assumption of equivalent variance-covariance matrices was not met although MANOVA procedures are robust to violations of this assumption. However, because smaller sample sizes produced larger variances and covariances, Pillai's Trace was used to evaluate the multivariate significance (Olson, 1976; Tabachnick & Fidell, 2001). While assumptions were not fully met for the parametric procedure, the less powerful nonparametric procedure produced similar results.

Following the MANOVA, descriptive statistics of mean and standard deviation were used to show scaled score differences and post hoc *t* tests compared which values of race and community type contributed most to the explanation of ELA and mathematics scores. To further examine differences, scores were coded into pass/fail categories. The coding

allowed comparisons of Black and White students for consequences of a particular score, passing or failing the test. In addition, scores were coded by previously given achievement levels to make comparisons.

As defined in the delimitations section of Chapter One, scores of students in alternative schools, charter schools, and university lab schools were not included in the analyses of scaled score differences, pass/fail differences and achievement level differences for reasons given previously. Scores of students in magnet schools were included in the quantitative analyses because magnet schools are a part of the reorganization of urban schools (Blank & Archbald, 1992). In the state of Louisiana, 81% of magnet schools with eighth graders are located in urban areas. An examination of the manner scores of students in magnet schools affected the results was needed to further explore how community type and race affect the achievement of Black and White students. Therefore, an analysis of LEAP 21 pass/fail rates was conducted to compare results when magnet school student scores were included versus when they were not.

Phase I allowed exploration of statistical differences between test performance of Black students and White students in various community types. The disparate effect high stakes testing places on Black students is also examined qualitatively by interviewing parents. The qualitative part of the study is described in Phase II.

Phase II

Gall, Borg, and Gall (2003) describe one use of qualitative research as a discovery role, contrasting it with a confirmatory role of quantitative research. Phase II used discovery for gathering parent perspectives on the impact of high stakes tests on their children. First, parents were asked what they understood about why students were tested.

In other words, did parents understand that testing was part of educational reform and part of raising standards in our schools? Parents were questioned about the effect passing or failing the test impacted their children's learning. Finally, parents were asked about the fairness of laws and regulations governing students to whom high stakes consequences applied. That is, did parents believe these consequences applied equally to students who were rich or poor, Black or White, male or female, and public or non-public? Other fairness issues included use of test results as the criterion for promotion to ninth grade, and administering the test as students complete grade 8. Because opinions may differ by the race and community type of parents' children, perspectives were gathered and analyzed by these categories. Five research questions guided Phase II:

1. What are the perspectives of parents of eighth graders about why students take LEAP 21? Parent perspectives were explored for both meeting the reform intent of raising standards in schools and increasing student learning by use of high stakes consequences.
2. What are the perspectives of parents of eighth graders about the fairness of the test on groups of students? Test fairness was explored for impacting groups of students who are rich or poor, Black or White, male or female, and public or non-public.
3. What are the perspectives of parents of eighth graders about using the test as a single criterion for promotion and timing the test as students complete grade 8?
4. Do perspectives of parents of eighth graders who pass and fail LEAP 21 vary based on race, Black and White, and community type; rural, suburban, and urban?

5. Do perspectives of parents of eighth graders who take LEAP 21 suggest a disparate effect on poor and minority students in Louisiana?

The current study used parent focus groups to gather perspectives on whether Louisiana mirrors the rest of the country in producing test score disparities along racial lines.

Sampling

Suitable districts and schools were identified before parent sampling took place. Districts were selected purposively (Tashakorri & Teddlie, 1998) from the southern region of the state to control for regional differences. Districts contained schools that met the criteria (Creswell, 1998) of having students who tested as eighth graders in the three community types. District superintendents and school principals granted permission for parents to be selected. Appendix A provides the district and school permission letters. Six schools were used because of low numbers by a particular race for parents of students who failed the test.

Parent sampling took place based on the placement of the student in the 2003-04 academic year as a result of passing or failing LEAP 21. The sample consisted of parents of students who passed the high stakes tests and were currently ninth graders. Also, parents of students who failed the test and were in grade 8.5 participated. In other words, parents were selected based on their children meeting all high stakes test requirements, a ninth grade placement, or lacking the high stakes test requirement, a grade 8.5 placement. Parents of students who were repeating eighth grade were not included as large percentages of repeaters are special education students, a variable beyond the scope of

this study. In a poll of one southern district, the percentage was found to be 73%, thus making the exclusion suitable.

Focus group sampling was homogeneous for race to facilitate self-disclosure (Jourard, 1964; Krueger, 1994). In addition, groups were homogeneous for parents whose children passed or failed the test, except in the suburban community type. While school liaisons of a counselor and secretary assisted with parent recruitment in rural and urban community types respectively, the district liaison in the suburban community type, a supervisor, allowed parent recruitment through principals on parent report card day. Even with further requests and follow-up phone calls, anticipated participation remained low. Therefore, only two focus groups, which were homogeneous by race and heterogeneous to passing or failing the test, were held in the suburban community. Table 5 details the focus group sampling. In rural and urban community types, parents were recruited

Table 5

Parent Focus Group Sampling

Community Type	Passed		Failed	
	Parents of Black Students	Parents of White Students	Parents of Black Students	Parents of White Students
Rural	Group 1	Group 2	Group 3	Group 4
Suburban	^a Group 5	^b Group 6	^a Group 7	^b Group 8
Urban	Group 9	Group 10	Group 11	Group 12

^a Parent perspectives were gathered within the same session.

^b Parent perspectives were gathered within the same session.

from a printout of students by grade and race. Focus groups in the three community types were heterogeneous for variables of socioeconomic status, gender, and one or two parents

in the household. A token of appreciation, a \$2 bill, was given to students who had their parents complete participation forms. Follow-up phone calls were made to ensure a strong turnout. To encourage parent participation a \$20 bill or a restaurant food token (Krueger, 1994; Morgan, 1997) was provided. An assistant moderator (Krueger, 2004a), with research methods coursework, was used to help with refreshments, welcome arriving participants, and debrief with moderator

. The groups comprised cases for the exploration of parental perspectives. The study followed a Yin (1994) Type 3 design, or “holistic multiple cases” (Yin, 1994, p. 39). In this phase of the study, parents were the unit of analysis and multiple cases were used. The next section describes the interview instrument.

Instrumentation

Marshall and Rossman (1999) explain the purpose of a focus group interview as “asking focused questions to encourage discussion and expression of differing points of view” (p. 114). The interviews allowed a variety of perspectives to be expressed in a limited amount of time (Krueger, 1988; Marshall & Rossman, 1999) by using different types of questions (Krueger 2004a). An introductory script was read followed by a round robin opening question. Seven key questions followed (Krueger, 2004a) including a “think back” question, taking parents specifically back to the eighth grade testing experience. The questions explored reform intent in asking parents why students take LEAP 21. Other questions explored the impact on student learning asking, “How has LEAP 21 affected what your child is taught at school?” and “When a student passes the test, what do you think that says about the student’s learning?” Parents were asked about the test as a single measure for promotion and timing the test “right before your child

finished eighth grade.” Because testing may advantage some groups of students over others, parents were asked about LEAP 21 impacting students who are rich or poor, Black or White, and male or female. Fairness was also used in asking parents about nonpublic schools being exempt from high stakes consequences. The session ended with a question allowing parents to reflect on the entire discussion (Krueger, 2004a) by asking, “Of all things that we have talked about, what do you think is the most important?”

Appendix B provides the instrument used in the study.

The instrument was piloted on parents of Black grade 8.5 students in April 2004. Because of piloting the instrument, one phrase, “because of the LEAP test”, was added to the question about considering nonpublic school for the parent’s child. Other questions remained the same. Besides practice in interview instrument use, the pilot also allowed the researcher to become familiar with an assistant moderator, the tape recorders, the site set-up and refreshments, and the time needed for conducting a focus group session.

Data Collection and Analysis

The Systematic Analysis Process (Krueger, 2004b) was selected for focus group data collection and analysis. The process allowed listening for inconsistencies and probing for understanding with questions such as, “Would you explain further?” and “Would you give an example?” while still in the sessions and before analyzing for themes. Two recorders were used in gathering parent comments as limited notes were taken.

Immediately following the interviews, a diagram was drawn of the seating arrangement. Debriefing took place between the moderator and assistant moderator. Hunches, themes, and interpretations were written, as were notes on the manner the group compared to other groups. Appendix C provides the form to record post focus group notes.

After each session, the data were transcribed and a report was made to summarize the data. As multiple sessions were conducted, the data were analyzed for emerging themes, which came from the participants' words but also from the context, consistency, frequency, intensity, and specificity (Krueger, 2004b). Pattern matching (Yin, 1994) did not take place as parent perspectives were largely missing on high stakes assessment. However, through the focus group instrument, questions were asked about three themes of reform intent, student learning, and test fairness; and, additional themes emerged from the data. The data were analyzed by categories of parents to allow understanding of the extent race of the student, passing or failing the test, and community type influenced parent perspectives. Stake (1995) recognizes the role of the researcher as a factor in data analysis and presentation. In the current study, the role moved from biographer of the reform in the literature review, to teacher in presentation of key factors affecting testing, and finally interpreter through data analysis (Stake, 1995). As to the manner trustworthiness was ensured, a description takes place in the next section.

Trustworthiness

Trustworthiness involved several measures. Because of the sampling scheme, reflexive journaling took place to record the reasoning of the researcher in making decisions for parent selection. Peer debriefing was used for study clarity. Thick description, or the detailed description of all information (Tashakkori and Teddlie, 1998), was used to provide evidence for findings and interpretations. The researcher used methods of inference quality to make good study decisions, but also to ensure replicability. Discussing trustworthiness completes the discussion of the qualitative phase of the study. The next section provides a summary of the methodology chapter.

Summary

Chapter Three included the methodology necessary to answer the research questions. The questions guided the mixed method study utilizing two phases. Each phase was described allowing readers to follow the manner the questions were answered. Phase I used a multivariate analysis of variance (MANOVA) to examine test score differences of LEAP 21 for Black and White students in rural, suburban, and urban schools. Pass/Fail rates and achievement level differences were also examined for scores of the students in the three community types. Magnet school student scores were examined to describe the differences when the student scores were included. Phase II used focus groups to gather parent perspectives of the impact of the tests on students. Parents were selected based on the grade placement of their child and also by their race and community type. The interview instrument and questions were described as well as the means of data collection and reporting. Methods of reflexive journaling, peer debriefing, and thick description were given to ensure the trustworthiness of the results of the study. The next chapter gives the results of the quantitative part of the study.

CHAPTER 4. QUANTITATIVE RESULTS

The study presents a quantitative and qualitative examination of high stakes testing of Louisiana eighth graders. The quantitative results are given in the current chapter and the qualitative results are given in Chapter Five. Chapter Four begins with an analysis of scores with high stakes consequences, English language arts (ELA) and mathematics LEAP 21 sub-tests, for Black and White students in three community types; rural, suburban, and urban. A multivariate analysis of variance (MANOVA) examines sub-test differences by race and community type with post hoc *t* tests revealing specific differences. Cohen's *d* provides effect sizes. Additional analyses explore the pass/fail rates and achievement level differences for the same variables. Finally, because an interaction between race and community is shown, an examination of how magnet school students impact the test data takes place.

Phase I

Data for the quantitative analysis were received from the Louisiana Department of Education (LDE). The data contained 232,329 scaled scores of students taking high stakes tests in the spring of the 2002-03 school year. Scores of students taking only one subtest were excluded, as were scores of students not in grade 8 and who were not Black or White as described by the delimitations of the study. Also, student scores came from public schools without special enrollment criteria of alternative, charter, or university lab school for reasons given previously. An additional exclusion took place. Student scores from new schools were excluded because the schools lacked assignment of a community type code.

Number and Location of Students

In all, 47,535 eighth graders took LEAP 21 and received scores in ELA and mathematics. Slightly more White students (52.11%) took the test than Black students (47.89%). In addition, in urban areas, almost twice as many Black students (22.80%) tested as in rural (13.1%) or suburban (11.96%) areas. In suburban and rural areas, White students were tested (21.64% and 22.59% respectively) about three times the rate of White students in urban (7.88%) areas. Table 6 provides the number and percentage of students in the state by race and community type used in the current study.

Table 6

Eighth Graders by Race and Community Type

Race	Rural		Suburban		Urban		Total	
	n	%	n	%	n	%	n	%
Black	6,241	13.13%	5,685	11.96%	10,837	22.80%	22,763	47.89%
White	10,737	22.59%	10,287	21.64%	3,748	7.88%	24,772	52.11%
Total	16,978	35.72%	15,972	33.60%	14,585	30.68%	47,535	100.00%

Location of Louisiana schools by community type and district is given in Figure 2 revealing the pre-dominantly rural setting for the state. Four districts contained all urban schools; Orleans Parish, East Baton Rouge Parish, Caddo Parish and the City of Monroe. Also urban schools were mixed with rural and suburban schools in Lafayette, Calcasieu, Rapides, Bossier, Ouachita, St. Tammany, and Terrebonne Parishes. Districts with pre-dominantly suburban schools surround the urban districts. Districts where only suburban schools were found included: St. Bernard, Jefferson, St. Charles, St. John, St. James, St. Landry, and Webster Parishes. Mixtures of suburban and rural schools were found in

West Baton Rouge, St. Martin, Iberia, Acadia, Ascension, Livingston, and LaFourche Parishes. Forty school districts contained only rural schools.

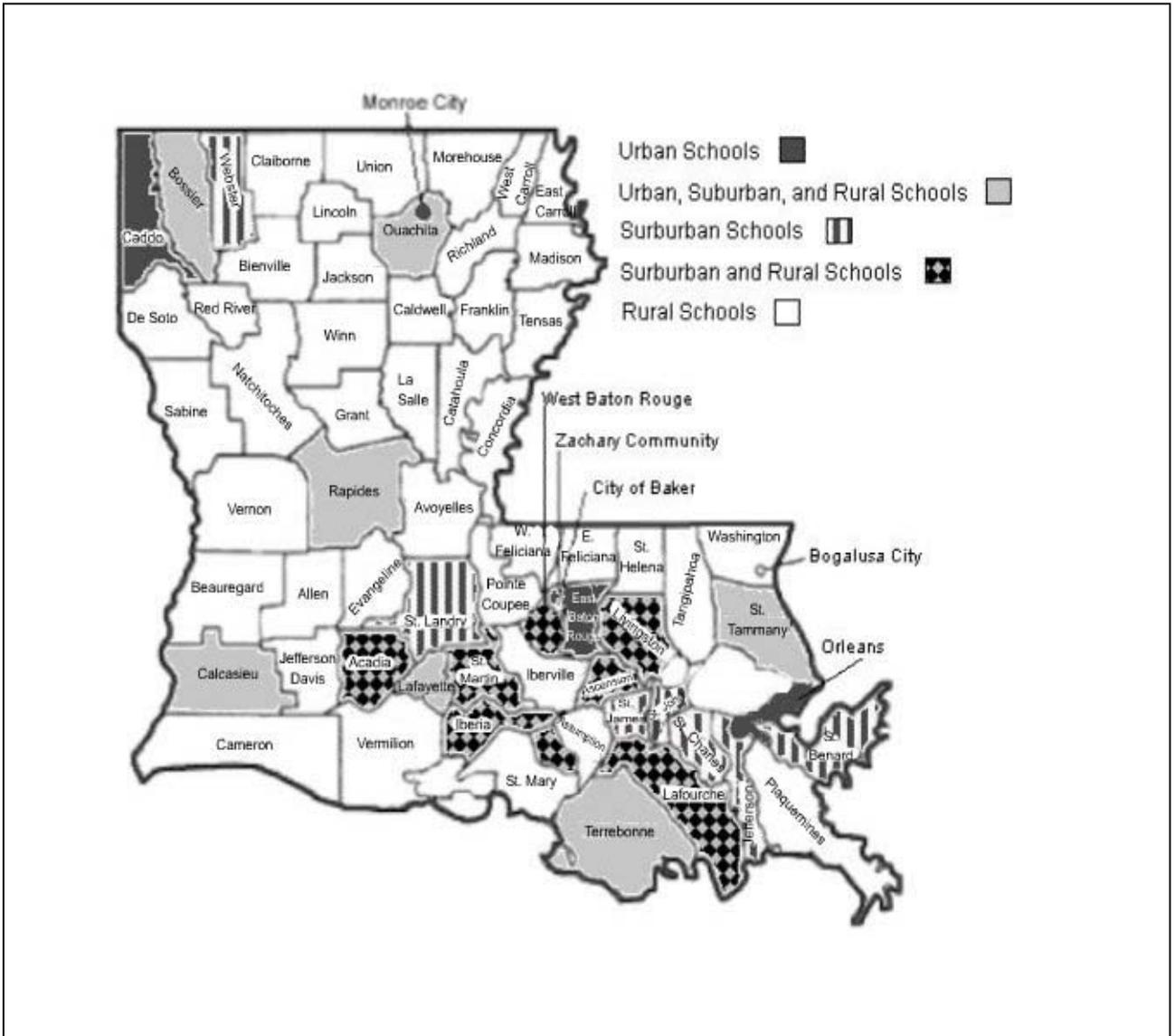


Figure 2

Louisiana District Map Showing Schools by Location

From the Louisiana Department of Education website, <http://www.louisianaschools.net/lde/bese/1636.html>, coded by community type

Race and Community Type

In examining the relationship between the independent variables of community type and race on the dependent variables of ELA and mathematics scaled scores, a multivariate analysis of variance (MANOVA) was used (see Table 7).

Table 7

Multivariate Analysis of Variance (MANOVA) Results

Variables	Pillai's Trace Value	F	Hypothesis df	Error df	Sig.
Community Type	.00	2.1	4	95,058	.07
Race	.18	5137.7	2	47,528	.00*
Community Type * Race	.01	131.6	4	95,058	.00*

* $p < .05$

Given a Pillai's Trace value of .01, there was a statistically significant interaction effect ($F=131.6$, $df=4,95058$, $p < .05$) for community type and race. While Black students scored similarly in rural and suburban settings, their scores dropped in the urban community type for ELA and mathematics. White student scores followed a different trend. While scoring similarly in rural and suburban settings, White student scores rose in both subtests in urban settings. The mean scaled scores by race and community type for the two subtests are given in Table 8.

Table 8

Mean Scaled Scores by Race and Community Type

		Community Type					
		Rural		Suburban		Urban	
Sub-test	Race	Mean	Std.Dev.	Mean	Std. Dev.	Mean	Std. Dev.
ELA	Black	298.78	40.58	299.77	41.85	290.76	45.08
ELA	White	329.23	36.28	329.76	36.57	340.10	37.58
Math	Black	300.11	46.19	300.24	46.00	290.96	50.79
Math	White	334.49	39.51	335.55	40.13	346.04	42.05

Figure 3 provides ELA scores by race and community type. The cutoff score for passing the test is also shown. The figure illustrates how the mean scores for Black and White students across all community types are found above the minimal ELA passing score of 269.

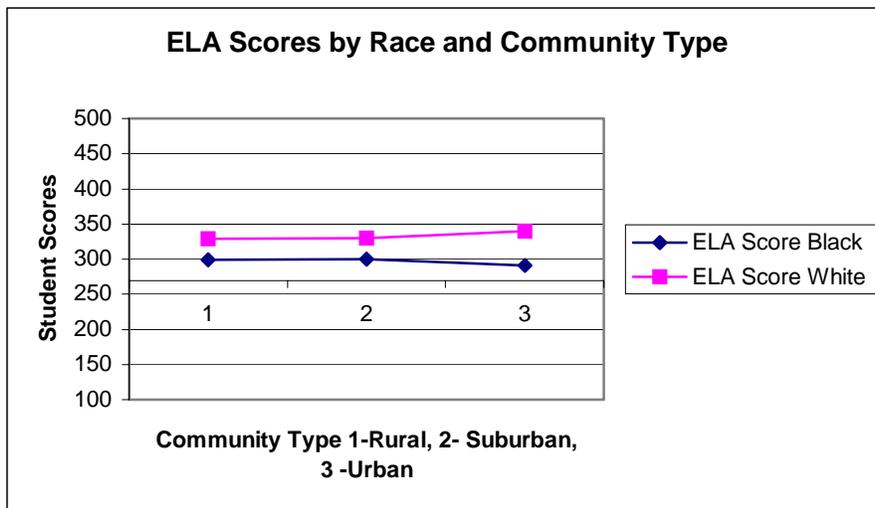


Figure 3

ELA Mean Scaled Scores by Race and Community Type

Figure 4 gives the mean mathematics scaled scores of Black and White students across the three community types along with the cutoff score for mathematics of 296. The figure is useful in showing how Black students meet the passing score in rural and suburban community types, but not in urban areas. At the same time, mean scores for White students for all community types are found at least 30 points above the cutoff score.

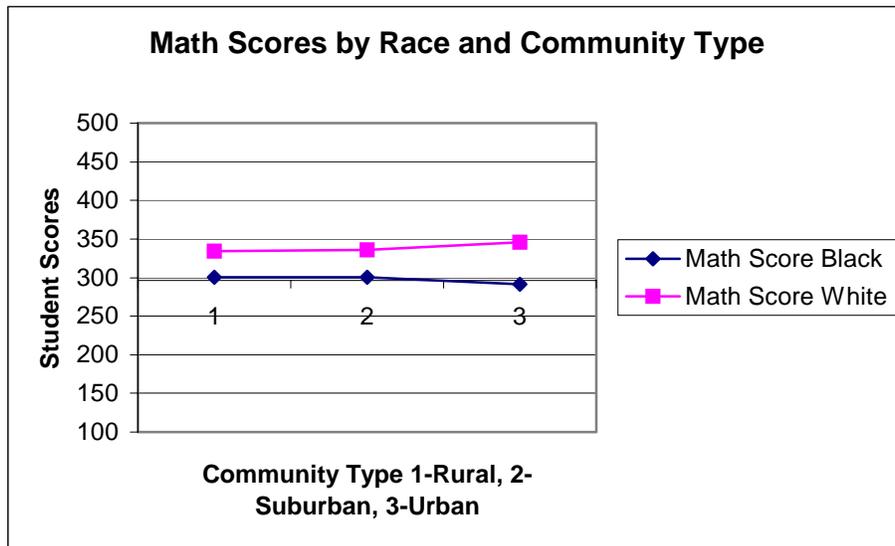


Figure 4

Mathematics Mean Scaled Scores by Race and Community Type

To examine specific differences in the interaction of race and community type, post hoc *t* tests were run. Mean scaled score differences by race ranged from 29.99 in the suburban community type for ELA to 55.07 for urban community type for mathematics. Equal variance was not assumed as the smaller *n* was paired with the larger variance (Hopkins, Glass, & Hopkins, 1987). Differences for the interaction of race and community type were significant across all categories. Because statistical significance could easily be predicted, reporting the effect size was needed; Thompson (1994)

supports reporting them to provide the reader a measure to evaluate the results.

Therefore, effect sizes were examined using Cohen’s *d* statistic. Medium effect sizes

(Cohen, 1988) were found for rural ELA and mathematics as well as suburban ELA.

Large effect sizes (Cohen, 1988) were found for suburban mathematics and urban ELA

and mathematics. In fact the largest effect size was found for urban ELA (see Table 9).

Table 9

t-tests and Effect Sizes for the Interaction of Race and Community Type

Community Type	Sub-Test	Mean Diff.	<i>t</i> test		Sig.	99% C. I.		Cohen’s <i>d</i>	Std. Error Diff.
				<i>df</i>		Lower	Upper		
Rural	ELA	30.45	48.99	11,893.62	.00	28.85	32.06	.78	.62
	Math	34.39	49.26	11,468.71	.00	32.59	36.19	.78	.70
Suburban	ELA	29.99	45.31	10,462.48	.00	28.29	31.70	.75	.66
	Math	35.31	48.57	10,449.23	.00	33.44	37.19	.80	.73
Urban	ELA	49.35	65.69	7,743.24	.00	47.41	51.28	1.24	.75
	Math	55.07	65.37	7,796.16	.00	52.90	57.24	1.23	.84

Community type was not shown to be statistically significant (see Table 10).

Student mean scaled scores across community type varied minimally, 1.4 points for ELA

Table 10

Mean ELA and Mathematics Scores Across Community Types

Community Type	ELA Mean Score	Mathematics Mean Score
Rural	314.01	317.30
Suburban	314.76	317.90
Urban	315.43	318.50

and 1.2 points for mathematics. Urban students scored slightly higher in both ELA and mathematics than their suburban and rural counterparts; however, this difference was not statistically significant.

Race was shown to be statistically significant at $p < .05$ (Pillai's Trace = .18, $F = 5137.7$, $df = 2,47528$, $p < .05$). Mean differences in ELA and mathematics for scores of Black students and White students are given in Table 11. The mean difference by race is greater in mathematics (41.6 points) than in ELA (36.6 points).

Table 11

Mean ELA and Mathematics Scores by Race

Race	ELA Mean Score	Mathematics Mean Score
Black Students	296.43	297.10
White Students	333.03	338.70

Pass/Fail Categories

To further examine differences in ELA and mathematics performance, scaled scores were coded into pass/fail categories. Category definitions followed coding by the Louisiana Department of Education (LDE) with a cutoff score of 269 for ELA, and 296 for mathematics (LDE Interpretive Guide 2001, 2003). For Black students in rural or suburban community types, one in five students (19.2%) failed the ELA subtest. For Black students in urban community types, more than one in four (28.6%) students failed ELA. On the other hand, about 1 in 20 White students (5.3% and 5.8%) failed ELA in rural and suburban community types. Further, less than 1 in 28 (3.7%) White students failed ELA in the urban community type. Pass/fail rates in ELA for Black and White

students across community types are provided in Table 12. Black students failed ELA at a rate 3-8 times that of White students in similar community types with the greatest difference in urban community types.

Table 12

Passage Rates in ELA Across Race and Community Types

Student Race	Rural		Suburban		Urban		Total	
	Count	%	Count	%	Count	%	Count	%
Black	6,241	100.0	5,685	100.0	10,837	100.0	22,763	100.0
Pass	5,043	80.8	4,596	80.8	7,740	71.4	17,379	76.3
Fail	1,198	19.2	1,089	19.2	3,097	28.6	5,384	23.7
White	10,737	100.0	10,287	100.0	3,748	100.0	24,772	100.0
Pass	10,169	94.7	9,693	94.2	3,609	96.3	23,471	94.7
Fail	568	5.3	594	5.8	139	3.7	1,301	5.3

For mathematics, similar results of Black students failing at higher rates than White students were found. For rural and suburban community types, more than one in three Black students (37.9% and 38.9%) failed the test. In urban community types, almost one in two (48.8%) Black students failed the test. For White students in rural and suburban community types, about one in eight students failed the test (11.8% and 12.3% respectively), whereas in urban community types about 1 in 12 (8.3%) students failed the test. The pass/fail mathematics comparisons for Black and White students across community types are given in Table 13. Black students failed mathematics at a rate 3-6 times that of White students in similar community types with the greatest difference in urban community types.

Table 13

Passage Rates in Mathematics Across Race and Community Type

Student Race	Rural		Suburban		Urban		Total	
	Count	%	Count	%	Count	%	Count	%
Black	6,241	100.0	5,685	100.0	10,837	100.0	22,763	100.0
Pass	3,876	62.1	3,476	61.1	5,548	51.2	12,900	56.7
Fail	2,365	37.9	2,209	38.9	5,289	48.8	9,863	43.3
White	10,737	100.0	10,287	100.0	3,748	100.0	24,772	100.0
Pass	9,470	88.2	9,022	87.7	3,436	91.7	21,928	88.5
Fail	1,267	11.8	1,265	12.3	312	8.3	2,844	11.5

Achievement Levels

Further analysis of the data by proficiency levels was made. The Louisiana Department of Education (LDE) categorizes scores by achievement levels (see Table 14).

Table 14

Achievement Levels for Grade 8 LEAP 21

LEAP 21 Achievement Levels – Grade 8		
	English Language Arts	Mathematics
Unsatisfactory	100-268	100-295
Approaching Basic	269-314	296-320
Basic	315-355	321-375
Mastery (Proficient)	356-401	376-397
Advanced	402-500	398-500

As to how Black and White students scored by ELA achievement levels, the largest percentage of White students scored in the Basic achievement level across all community types. The largest percentage of Black students scored in the Approaching Basic category across all community types. In effect, the largest percentage of Black students scored 1 achievement level behind White students in all community types. In addition, White students scored in the Mastery and Advanced achievement levels around 4 times the rate of Black students in rural and suburban community types. In the urban community type, the rate was greater than 6 times that of Black students in the Mastery achievement level and 14 times greater in the Advanced level (see Table 15).

Table 15

ELA Achievement Levels by Race and Community Type

ELA	Rural		Suburban		Urban		Overall	
	Black	White	Black	White	Black	White	Black	White
Unsatisfactory	19.2%	5.3%	19.2%	5.8%	28.6%	3.7%	23.7%	5.3%
Approaching Basic	45.4%	25.4%	43.4%	24.5%	41.8%	17.8%	43.2%	23.9%
Basic	29.8%	47.6%	31.7%	46.4%	24.6%	43.8%	27.8%	46.5%
Mastery	5.4%	20.7%	5.5%	22.3%	4.9%	31.8%	5.2%	23.0%
Advanced	.2%	1.1%	.2%	1.0%	.2%	2.9%	.2%	1.3%

In the content area of mathematics, the largest percentage of White students scored in the Basic achievement level across all community types. The largest percentage of Black students scored in the Unsatisfactory level across all three community types, ultimately failing to meet the cutoff score. Consequently, the largest percentage of scores of Black students was found 2 achievement levels below the largest percentage of scores of White students in all community types. White students scored in the Mastery and Advanced

achievement levels at rates greater than 5 times the rates of Black students in rural and suburban community types. In urban areas, White students scored in the Mastery level 8 times that of Black students and in the Advanced level 15 times the rate of Black students (see Table 16).

Table 16

Mathematics Achievement Levels by Race and Community Type

Mathematics	Rural		Suburban		Urban		Overall	
	Black	White	Black	White	Black	White	Black	White
Unsatisfactory	37.9%	11.8%	38.9%	12.3%	48.8%	8.3%	43.3%	11.5%
Approaching Basic	28.9%	19.7%	27.8%	19.2%	24.1%	15.6%	26.4%	18.9%
Basic	31.3%	56.6%	31.1%	55.3%	24.9%	54.2%	28.2%	55.7%
Mastery	1.3%	8.1%	1.5%	8.2%	1.6%	12.8%	1.5%	8.9%
Advanced	.6%	3.8%	.8%	5.0%	.6%	9.0%	.6%	5.1%

Magnet School Impact

To gain further insight into the effects of race and community type on student performances on grade 8 LEAP 21, magnet schools were examined. Magnet schools are most often organized around a specialized curricular theme or subject matter (Blank & Archbald, 1992; Goldring & Smrekar, 2002). Louisiana had 68 schools classified as race balancing magnet, academic magnet, or both for the 2002-2003 academic school year. Race balancing magnet schools offer special programs for the purpose of attracting students to enroll voluntarily, thereby advancing school desegregation. For example, a magnet school may use technology, liberal arts, creative or performing arts to attract student applicants. An academic magnet, conversely, requires students to meet academic

entry criteria. Admission policies for the two types of schools require verification of eligibility and when demand exceeds capacity, lotteries occur.

Twenty-six magnet schools in the state conducted eighth grade testing in the spring of 2003. Twenty-two of these schools were race balancing magnet schools. The four exceptions were three schools in Orleans Parish and one rural school. The rural school used both race balancing and academic magnet criteria in its magnet program. Of the 26 schools, 21 were urban, 4 were suburban, and 1 was rural. The schools represented 2,757 eighth grade students who tested in the spring of 2003.

Comparisons of passing rates by Black and White students in the three community types were made when magnet school student scores were included and excluded (see Table 17). Rural and suburban passing rates for Black and White students in ELA and

Table 17

Passage Rates With and Without Magnet School Student Scores

Passage Rate % ELA						
Community Type	Black Students			White Students		
	Without	With	Gain	Without	With	Gain
Rural	80.8	80.8	+. 0	94.7	94.7	+. 0
Suburban	80.8	80.8	+. 0	94.2	94.2	+. 0
Urban	68.6	71.4	+2.8	96.0	96.3	+. 3
Passage Rate % Mathematics						
Rural	62.0	62.1	+. 1	88.2	88.2	+. 0
Suburban	60.9	61.1	+. 2	87.6	87.7	+. 1
Urban	48.5	51.2	+2.7	91.2	91.7	+. 5

mathematics remained within .2 percentage points across all categories reflecting the small numbers of magnet school students in rural and suburban school communities. However, when urban scores were examined, several differences were noted. Black students passed ELA at a rate 2.8 percentage points higher when magnet school student scores were included. The rate for mathematics was 2.7 percentage points higher. Whereas, when magnet school student scores were included, White student passage rates increased in ELA and mathematics by only .3 and .5 percentage points, respectively. Consequently, the pass rates for White students remained approximately the same whether or not magnet school passage rates were included; however, the pass rates for Black students improved in urban areas at a rate 5-9 times the rate of improvement of White students. Goldring and Smrekar (2002) recognize that districts with magnet schools on average invest 10% more money for instructional resources and almost three-fourths of the districts with magnet schools allot additional teaching staff for magnet schools. These two researchers also recognize the bottom-up reform of the magnet schools. As Goldring and Smrekar explain, teachers in magnet schools are empowered to teach and principals are empowered to transform these schools with minimal interference from top-down mandates. These findings contribute to policy considerations which are discussed in Chapter Six.

Summary

The quantitative results of the study were presented in the current chapter. Students were spread across community types allowing comparisons of students by race in the three school locations. A multivariate analysis of variance (MANOVA) was used to describe test score differences in ELA and mathematics of Black and White students. Race and the interaction of race and community type were statistically significant. As

could be predicted, post hoc *t* tests revealed all differences were significant. Medium and large effect sizes were found revealing the magnitude of the differences. An analysis of student scores by pass/fail categories revealed Black students failed the tests at rates higher than White students in similar communities, 3-7 times in ELA and 3-6 times in mathematics. The greatest differences were in urban school communities where nearly 1 in 2 Black students failed mathematics whereas, about 1 in 12 White students failed. Coding of scores by achievement levels revealed the largest percentage of White students was one achievement level higher than that of Black students in ELA and two achievement levels higher in mathematics. White students scored in the Mastery and Advanced levels at rates 4-14 times that of Black students in ELA and 5-15 times that of Black students in mathematics. When magnet schools were included, Black students in urban community types increased the percentage of students passing ELA and mathematics at a rate higher than the rate of increase of White students.

To fully describe the impact of high stakes testing, a qualitative analysis was needed. The results of the qualitative phase of the study are given in the next chapter.

CHAPTER 5. QUALITATIVE RESULTS

This chapter presents results from the second phase of the study, which used focus group interviews to gather parent perspectives. The chapter begins with sampling for district, school, and parent selection. Next, the method of data analysis and reporting are given. Four focus group reports follow allowing insights into parent conversations about testing (Krueger, 2004b). Next, the results are diagrammed by themes (Krueger, 1994, 2004b) across sessions revealing relationships (Krueger, 1994) within the perspectives and showing differences in opinions of parents by race of their children and community type of their child's school (Krueger, 2004b).

Sampling

Phase II began with district selection from one region of the state, the southern region. Purposeful district sampling allowed selection of multiple schools of the needed community types while controlling for regional differences. More than one school was needed to get 6-10 (Morgan, 1997) parents of Black and White students who passed and failed the test. In fact, three rural schools and two suburban schools were included because of small numbers of parents of Black and White children that failed LEAP 21. One urban school was needed for sampling of parents by categories. Also, to allow parents to give perspectives on the use of non-public schools to avoid the test consequences, each selected school was located within 2-10 miles from a non-public school. Sampling for the study and the number of parents who attended each focus group are given in Table 18. Parent selection from rural schools was random. Parents attended focus group sessions on two different nights at one of the schools; Black parents attended on one night and White parents on the other (Krueger, 1994). Each night, two sessions

were held allowing parents of children who passed or failed the test to give their opinions separately.

Table 18

Focus Group Sampling

Region	School Community Type	Number of Schools	Parents in Focus Groups	
			Black	White
Southern	Rural	3	5 - Pass 7 - Fail	5 - Pass 4 - Fail
Southern	Suburban	2	*(2 - Pass 3 - Fail)	*(1 - Pass 3 - Fail)
Southern	Urban	1	2 - Pass 10 - Fail	7 - Pass 4 - Fail

*Only two focus group sessions were held by race in the suburban community type.

Parents of students in suburban community types were selected from parents attending report card day in the district. Chapter Three described the manner parents were recruited through supervisor and principal liaisons. Because the two schools were located within pre-dominantly Black or White neighborhoods, the focus group interviews took place in a quiet room of a restaurant located mid-way between the schools. Participation rates were low for the four categories of parents and may be attributed to the logistics of utilizing schools within a geographically large district. Alternatively, the use of the supervisory liaisons may have contributed to low participation. The district supervisor controlled

access to parents through the report card day recruitment. Initially, 40 slips were distributed with 11 forms completed and returned. Follow-up phone calls were made to parents who did not return permission slips. Because numbers remained low, only two focus group sessions were held. One session contained all Black parents and the other all White parents. However, parents in the groups had students who both passed and failed the test and were seated accordingly. The moderator allowed one group to speak at a time. Krueger and Casey (2000) recognize the moderator's role in building an atmosphere of trust allowing participants to feel comfortable in sharing their perspectives. At the beginning of the interview, parents whose children failed the test were cautious about their responses; however, once they heard concerns of parents of children that passed the test, the dialogue became relaxed and flowed smoothly. The parents of students who failed the test no longer perceived themselves to be different from parents of children that passed (Morgan, 1997). Because of the seating arrangement, parent responses were transcribed based on their child passing or failing the test. Therefore, the focus group sessions were considered useful and contributed to the study, but are limited by the sampling procedure. More discussion is given in Chapter Six on the usefulness of the suburban sampling.

One school was used for recruiting parents for urban focus groups. Eligible parents were identified through a printout of students by grade and race. Selection from the printout was random with follow-up phone calls made. Because the quantitative analysis revealed a large percentage of Black students failed the test in urban community types, questions "What could you tell me about English/Language Arts and mathematics LEAP

tests? “and “Do students have a more difficult time with a particular subject?” were added.

In all, 53 parents participated in tape-recorded focus group sessions. One hundred eighty three pages of transcribed data were analyzed using the Systematic Analysis Process (Krueger, 2004b). The method involved probing for understanding while still in the group, summarizing each session into a report, and analyzing for themes across groups (Krueger, 1994). The next section provides reports for four of the interview conversations.

Focus Group Parent Reports

Four focus group session reports provide contrasts in parent perspectives by race, passing or failing the test, and community type. The reports summarize the sessions and allow the reader into the conversations of parents (see Table 19). The terms “Country Road” and “City Street” were used figuratively for parents of rural and urban students, respectively. Parents of suburban students, “Bypass Loop Highway,” were not included in the reports because of the limitations of the sampling procedure described previously.

The sessions were reported in three areas (a) reform intent - Why were students tested? (b) student learning – What were students learning? and (c) test fairness – Was the test fair? The reports, selected because of the number of participants and contrasts by categories of parents, are Country Road, Black, ninth grade parents; City Street, White, ninth grade parents; City Street, Black, grade 8.5 parents; and City Street, White, grade 8.5 parents. In these reports parents are referred to by a first name selected as a pseudonym to maintain parent confidentiality.

Table 19

Four Focus Group Reports

(26) Parent Conversations				
	(5) Country Road, Black, Ninth Grade	(7) City Street, White, Ninth Grade	(10) City Street, Black, 8.5 Grade	(4) City Street, White, 8.5 Grade
Why were students tested?	(3) The test was used for student promotion. (3) Parents asked why the test was given only in Louisiana. (2) Parents said they did not know.	(2) LEAP 21 determined a child’s academic level and graded the teacher. (1) The test had a hidden meaning.	(1) LEAP 21 showed the levels of students. (6) Parents said they did not know.	(4) LEAP 21 showed students could read: it prevented social promotion. (1) Students were tested to keep them from quitting school, but they quit anyway. (2) The test was used to bring students to the national average.

(table con’d.)

What were students learning?	(1) The content was difficult and students needed encouragement. (3) Capable students were “freezing up” on the test. (2) The test preparation was not important because skill sheets were sent home on “runoff paper.”	(2) Students learned more content to pass the test. (3) The Louisiana culture is “laid back” preventing the state from being compared in learning with others. (4) Passing the test meant students learned and the teacher taught. (4) The mathematics content differed from what parents were taught. (7) The pace and stress of the test kept students from learning.	(9) Learning did not increase because students “give up.” (6) Students were not taught the tested content. (3) Students needed early preparation to learn the content. (4) Students did not learn because teachers did not know the content.	(2) Students were learning more because of LEAP 21. (2) Failing the test discouraged students from learning by focusing on their weaknesses. (1) Students needed to be grouped by ability; otherwise, the curriculum was “dumbed down.”
------------------------------	---	---	--	---

(table con’d.)

<p>Was the test fair?</p>	<p>(2) Rich parents provided private tutoring for their children. (3) Race was not a factor in passing the test. (2) Girls did better on the test. (3) Parents said non-public schools were used to avoid LEAP 21. (3) The stress of taking LEAP 21 affected students.</p>	<p>(3) The test was no different for students in rich or poor families and Black or White families. (5) Use of the test as a single criterion for promotion was not supported. (3) Testing affects students physically. (3) Students were notified of the test results too late in the school year.</p>	<p>(7) White students in wealthy families had an advantage. (6) Politicians, who passed laws requiring public school students to test, sent their children to non-public schools. (10) Use of the test as a single criterion for promotion was not supported. (2) Parent expectations were reduced for their children because of failing the test.</p>	<p>(3) Parental support and inner motivation made a difference and not if a family was wealthy. (1) A parent supported the test as a single criterion for promotion. (2) LEAP 21 allowed students to “go on” to grade 8.5. (3) Failing to require non-public school students to pass a test was unfair.</p>
---------------------------	--	---	--	---

Note. The numbers in parentheses refer to parents in each group and number of parents who held each opinion. Parents, who indicated no response either verbally or by gesture, cannot be assumed to agree or disagree with other participants.

Country Road, Black, Ninth Grade

Five Country Road parents attended the session. Brenda, whose child experienced no difficulty with the test, came because she was a substitute teacher and wanted to share her experiences. She was more outspoken than the others. Her experiences reflected a parent perspective, but also, those of a substitute teacher who worked to prepare students for LEAP 21 testing.

Why Were Students Tested? Two parents discussed their experiences with standardized testing when they were students in school and agreed that they “don’t have an understanding of it (LEAP 21).” Previously, testing was used to tell parents about the progress of their children by reporting a child’s strengths and weaknesses. However, three, Country Road parents expressed that LEAP 21 did not show a student’s strengths and weaknesses, but rather “if you pass or fail.” Thus, for the three parents, students were tested to identify those who were promoted and those who were not.

What Were Students Learning? Three parents said that students who were capable of passing the test often failed it. Students focused on the consequences of the test rather than just taking it. However, when students passed, the three parents agreed, “they (the students) comprehended everything they were taught.” Two parents were more skeptical and added, “its just pot luck” when students received passing test scores. One parent expressed that students, who made passing grades but failed the test, felt as if they had cheated their way through courses in school. High academic achieving students did not pass the test; therefore, those students must have cheated in classes in school. Two other parents stated that LEAP 21 study material was sent home on “runoff paper” and that meant the content was not important. Trisha suggested putting “it (the study material) in

book form so that materials can go back and forth from school to home.” Brenda stated that the hardest content in preparing students for the test was the measurement section. She had to get her son-in-law to help her. Also, as a substitute teacher, Brenda stated it was her job to “uproot all the negative” feelings students had about the test by making them realize that they could understand the content, even when students were discouraged by the difficult content.

Was the Test Fair? Two parents shared that rich parents possessed an advantage in being able to provide private tutoring. They contrasted private tutoring of a 1:1 teacher-student ratio with the 1:10 after school tutoring and suggested private tutoring helped students learn more. Three parents agreed that race was not a factor; and Trisha shared, “I figured the percentage (by race) is about the same with passing and failing.” Two parents expressed that girls in general did better on the test. “They (the girls) know a lot of stuff, more than the boys.”

As to the fairness of students attending non-public schools to avoid LEAP 21, three parents expressed “that’s what it is going to come to.” Country Road parents were aware of parents using non-public schools to avoid testing. Daria was ready to transfer her child to a non-public school because of a prior experience with an older child. The older child had been denied a high school diploma because she had failed the exit exam. The parent stated, “I am not doing it this time. I will not watch my child cry and think she’s nothing because of LEAP.” Parents agreed that sacrifices would have to be made for the money to be available for the non-public school tuition. Things like a cable bill would “go undone” to afford tuition. In fact, parents discussed non-public schools when discussing a high school diploma, even though parents had not yet been asked about the schools. However,

Brenda had not considered non-public schools because her son had no difficulty with “passing LEAP.”

Three, Country Road parents talked about the emotional impact of failing LEAP 21 on students. Trisha said her child asked, “If I don’t pass this test, are you still going to love me?” Another parent expressed that students would not be able to “take failing the test” explaining that “a tragedy” would take place because of the pressure. Two parents gave losing weight as an example of the way students handled the stress.” You know my daughter; she was never a smart child. She lost a lot of weight because she was stressed.” However, Brenda’s daughter became nervous about the test “the week-end prior to taking it (LEAP 21).”

In summary, parents had limited knowledge of why students were tested and shared that a student who could pass courses often failed LEAP 21. Also, parents believed that the wealthy possessed an advantage for helping their children and that a non-public school was a means of avoiding LEAP 21 consequences. Finally, student stress levels varied from losing weight to worrying about the test the weekend before students were tested.

City Street, White, Ninth Grade

Seven City Street parents attended the interview session. The perspective of Kyle, reflects the perspective of a parent whose child passed the test after going to summer remediation.

Why Were Students Tested? Two parents shared that LEAP 21 takes place to “determine a child’s level, and to grade the performance of the teacher.” Jane added that parents were “more or less” told by the school system of that reason for testing their

children. However, Kyle asked about the “whole story behind the test because all of a sudden it just popped up.” The questioning by Kyle suggested a hidden meaning for testing beyond knowing the level of a child and the extent that the teacher taught the content. Other parents were unable to answer his question.

What Were Students Learning? Two parents spoke about the press to pass the test as a cause for students learning more with LEAP 21. Renee’ shared that “teachers have to teach it (the curriculum) now because they want their students to know more in order to take the test and score better.” Mindy described the manner that students learn more. Teachers “cram a lot more studying in them (students) because they have to learn their regular curriculum plus do extra work toward the LEAP.” Also, according to Renee’ learning more was not “bad because our ratings of our children are really low compared to everywhere else.” However, three parents recognized “the way of life in Louisiana” is different, because the culture is more “laid back”, “less demanding” and “more friendly” than in other areas of the country. Therefore, being low in test scores compared to other states did not mean low in academic ability of the students, but rather, more about the culture of valuing a slower pace of living.

Renee’ admitted that her son” whizzed” through the test as it really wasn’t a problem for him.” However, four parents added that “some kids can take a test like that or they can’t.” When students passed the test, in the view of four parents, it meant students learned and that the teachers taught what students needed to know for the test. Four parents suggested the way mathematics was taught contributed to more students failing. Mary described that both she and her husband earned a high school diploma 19-20 years

ago; however, “the math has changed. It is too advanced.” Her older son had to help her eighth grader. “He (the older son) is good in math, so he helped her a lot with it.”

Kyle expressed that his daughter “was taught by good teachers, loves coming to school, and likes her teachers.” For Kyle, school represented building “character and personality. You make friends. You learn things. Then after all that, they give you a test. You fail this test, and you fail for the year.” Four parents realized that their children measured the school year by the test. Students knew that they could do poorly all year; however, they could still pass to the next grade by passing LEAP 21. Consequently, three parents agreed that the test created a “double standard.” Do poorly all year, pass the test, and you are promoted. Do well all year, fail the test, and you fail for the year. Kyle acknowledged that his daughter was not told “fail for the year, but pass the test” specifically, but she knew of the high regard for passing the test from the “pep rallies and all the big hype for the test”. In all, the seven parents felt the pace and stress caused by the test prevented the students from learning as they were capable of doing.

Was the Test Fair? As to the impact of the test on groups of students, three parents expressed no differences for rich or poor students, or Black or White students. When asked about males or females, Kyle said, “Girls cry when they fail.” Also, three parents described stress on students because of the test. Jane added, “Just the idea of it (the criterion to pass to ninth grade) being a test, they (the students) go crazy.” Kyle’s niece “threw up the whole week and she’s an honor roll (student). She did really good on the test, but just knowing the consequences of failing that test” caused the stress. Mary, whose child did not worry about the test, said that if the child failed he could not go to Boy Scout camp. Leona spoke about students not knowing if they should participate in

summer band camps. If the child fails the test again, the child should not have been in band camp; however, if the child passes, the child may not be eligible for band because of failing to attend the summer camp. In effect, the test negatively impacted the manner students participate in activities and prepare for the coming high school year. Three parents agreed that students are informed about test results late in the school year, especially for making summer plans. Kyle said, “You don’t know what school your kid is going to if they don’t pass at the end. It’s all up in the air for so long.” Kyle told of the delay in getting test results, which made buying the correct school uniform for his child difficult.

Kyle shared that he had no problem with non-public school students not being administered the test. He wished his “kid didn’t have to take it either.” Kyle knew he could not afford to send his daughter to a non-public school, although he did discuss the option with the junior high school counselor. Renée had previous experiences with sending her son to a non-public school. She said, “I pulled him out (of the non-public school), because I found out there were teachers that weren’t even certified.” Thus, the parent shared her belief that non-public schools compared poorly with public schools, but she had no testing measure for comparing the quality of the schools as public schools were compared. Renée further explained that she “believed in this test.” She just did not think “it should hold a child back (in a grade) if the child’s doing really good.” Four parents agreed and Corey summed up the group’s sentiment with “that’s not fair to these kids when they work so hard throughout the year.” Kyle agreed as his daughter was on the A/B honor roll; however, she started worrying about LEAP 21 as soon as she passed seventh grade.

In summary, City Street, White, ninth grade parents expressed that testing meant finding your child's academic level and they had limited understanding of why a high stakes test was needed. Parents agreed that students were learning more, because the test required that the students cover more content. On the other hand, parents questioned a hidden meaning behind testing without suggesting what it might be. Parents said the way mathematics was taught was a problem and expressed that school was more than taking a test. Parents shared feelings of stress that students have about testing. The stress meant physical responses and a concern about students reallocating time from summer activities to summer school. Lastly, parents explored the way non-public schools compared to public schools for teacher certification and not having to administer a test to students. However, parents agreed that the test should not be used as a single criterion in the public or non-public school systems.

City Street, Black, Grade 8.5

Ten City Street parents attended the focus group session. Parents in the group shared their opinions freely with the session lasting longer than the others, because of group size. The focus group was the largest of all sessions, indicating a high level of interest by parents in ways the test impacted their children. The high level of interest differed from the often-quoted criticism of Black, urban parents. In fact, one parent in the group addressed the stereotype of parent disinterest. "When they (the students) go to school, the first thing the teacher says is that it (the lack of learning) is the parents' fault. You cannot blame us, if you're bringing outside stuff (instructional content) into the school the kids don't understand and teachers don't understand."

Why Were Students Tested? Four parents stated that they did not know why students took the test. Two additional parents questioned from where the test came and for what purpose it came. One parent volunteered that testing took place to know the levels of the students and another parent agreed. When later speaking of test impacts, a parent added that if LEAP 21 was to stop unemployment, then the test was not being successful. The parent described students dropping out of school lacking the diploma needed to get a job “even in a fast food restaurant.” Other than the two comments, parents expressed no knowledge of why students were administered the test. Two parents referred to the test as “stupid” reflecting a level of frustration with the test.

What Were Students Learning? Nine parents agreed that student learning was not increasing because of the test. Rather, testing caused students to “give up on school.” One parent summarized what her child was feeling with “I (the student) took the test. They (the school system) want to hold me back. I can’t go to school all over again.” Six parents talked about instruction in the classroom being different from what was being tested on LEAP 21. Parents noted that regular instruction took place, but LEAP 21 asked questions in a manner “that went beyond the regular instruction”. The 10 parents said the gap was especially strong in mathematics. One parent commented that regular instruction included skills such as “ $1+1=2$.” However, questions on the test required a student to figure out word problems. Five parents questioned their ability to help their children when they did not understand what was being asked by test questions. Tara disliked questions asking which answer was “more right” than others and asked, “Why do you want me to think like you do? I don’t think like you.” Shayla described worksheets being sent home with the students as “Chinese” because she did not understand what was being asked or the

approach to help her child learn. Derek described a problem involving pies. “I have a large piece of pie. Terry has a larger piece. Dora has two pieces. Stephanie has three, but Dora’s two pieces might be larger than those three. I’m thinking in terms that you’ve got two pieces, so you’ve got more pie than mine.” Derek expressed that nowhere did they tell you the size of the pieces of pie. Derek asked that the mathematics questions be given in a manner where he could understand what was being asked. Melanie noted that when the test came for the student to take, the questions were re-worded causing students to “feel lost”, and to “give up.” The feelings further caused “the students to close that doggone (test) book and say, ‘Fail me.’”

Three parents said that LEAP 21 was introduced a couple of weeks before testing. The parents wanted their children to have more time to prepare and wanted the content taught all year in the manner students needed to meet test requirements. Also, three parents asked for a “strong foundation” in mathematics for their children and blamed the school system for not providing one. Marla said the child was robbed of “a common math in grade school” when not given the basic preparation needed for taking the test.

Four parents suggested that teachers do not know the mathematics content and then described the answer key to all the problems in the teacher’s edition of the textbook. While the teacher was able to give correct answers from the key, teachers lacked knowing the steps to teach the students to do the problem. The parents blamed teachers who were not trained in mathematics content as the reason. “You’ve got coaches teaching mathematics. English teachers teaching mathematics. P.E. teachers teaching mathematics.” In other words, teachers could not teach the content because of their lack of training, but they could give the correct answers by using the answer key in the

manual. Two parents blamed the way the school administration scheduled teachers' classes as the problem. Terri offered, "I heard a teacher say they (the administration) want her to teach three subjects in the class." Also, Sally shared that she had been encouraged to move to a different neighborhood for her child to have good teachers. She was told, "The better the neighborhood, the better the teachers." Sally expressed a willingness to move, but she lacked the money. Sally added, "You pay my way and I'll go."

Was the Test Fair? Ten parents said the test should not be used as a single criterion for promotion, "especially when they (students) make good grades all year long." If the test is what counts, then three parents suggested that school officials let students stay home for three or four years preparing for the test and come to school when ready to test. Students "don't need to go to school everyday if school is just about a test." Also, two parents made the same comparison when talking about teacher preparation in college. Derek said, "That's like someone that never went to college, fresh out of high school. They can take the NTE (National Teacher Exam) and pass it, cheating the school system." The parents suggested that officials require college students who wanted to be teachers only to pass teacher certification exams. College preparation was unnecessary if school was only about passing a test. Marla added that school had become taking a test as it "decided the child's life."

Four parents requested that teachers take the eighth grade test. If they fail, "suspend them and take their job," until they pass it. Parents also wanted to "select five random school board members and the superintendent" to take the test.

Seven parents agreed that being White meant you were “better off.” The description included politicians getting to make LEAP 21 decisions for public school students and having their own children in non-public schools. Derek quizzed the group with, “What religion are most politicians?” and “What race is predominantly Catholic?” Marla asked about the way a politician represented her concerns about testing for her child when the “politician’s own child” was not in public school. She summarized her feeling with “So, you (the politician) can throw me any kind of garbage in public school because your child is not even there to catch it.” Derek explained that the test itself did not address the Black culture. He said that if there was a segment on Black soul food, a Black ninth grader could pass, but others not exposed to the Black culture would fail. He suggested that the child knowing the Black culture would get “the job” and others would not. Marla further described that children hate their own Black color because they see other kids, “making it.” Marla said her child told her, “Mama, I’m tired of being Black. Because I’m Black, I can’t do this and I can’t do that.” Kenneth suggested for parents to “look when you bring your child to summer school. There’s a whole bunch of Blacks” and very few White students.

Two parents described stress as a factor for their children as students prepared for testing. However, different from any other groups, City Street, Black parents of grade 8.5 students conveyed high levels of frustration resulting from their children failing to meet parent expectations. Letha told of the expectations of parents for their children with, “I want diplomas on my wall.” However, the reality was students were not getting the credential, because they could not pass the test. Moreover, parents realized that passing the test was not needed for being successful. Even though expectations were diplomas,

eight parents stated that their children were considering lowered expectations, a Graduate Equivalency Diploma (GED) or quitting school. Marla said her son was telling her “he is not going back to school and I don’t want to hear my child saying something about killing himself because of a test he doesn’t understand. I’ll pull him out of school.” Even though parents expected high school diplomas, they were “pulling the child out” of school when the student failed the test and suggested physically harming himself or herself. Paula explained that she had a “son go through this school and he’s in the Navy now, and he didn’t have to go through all that (the test).”

In summary, City Street, Black parents lacked knowledge of why students were tested. The testing contributed to students “giving up on school” with mathematics more of a problem than English Language Arts (ELA). Parents found little connection between what students learned in the classroom and what was tested. Parents were aware that neighborhoods impacted the quality of teachers for their children. Parents lacked the ability to help their children prepare for testing and viewed the current school practices as inadequate preparation for the test. Also, parents elaborated that testing differed for White children whose parents were politicians and had their children in non-public schools. In all, parents expressed high levels of frustration caused by the testing requirements.

City Street, White, Grade 8.5

Four City Street parents attended the focus group session. Even though a married couple attended, their perspectives differed. Their discussion of the topics added to the interview, because they disagreed with each other. Thus, rather than one spouse being a

“silent partner” (Krueger, 1994, p.78), topics were sorted and clarified with less moderator questioning.

Why Were Students Tested? Four parents expressed their understanding of why testing took place. It meant accountability. For too long, “children that graduated got out into the world and couldn’t read.” Students were passed along from grade-to-grade being socially promoted, because they were too old. Also, two parents said that Louisiana schools needed to reach an academic national average like other states. Josie understood trying to reach the national average as she had lived in a northern state and her children were far ahead when she moved back to Louisiana. In fact, Josie suggested her children became bored in Louisiana schools.

Sarah suggested that testing took place to prevent students from dropping out by tracking their progress and making sure they were learning. But, she estimated that more students were quitting school because of LEAP 21. The students were “older by the time they got to their senior year, and they became discouraged.”

What Were Students Learning? Two parents said students had to “learn more stuff at a younger age.” Parents disagreed as to what “stuff” had to be learned. Tom thought too many social activities were a part of the curriculum “as politically correct movements” took place. He encouraged schools to stick to the “3 Rs.” Sarah and Josie suggested that students were learning more than they used to before LEAP 21. When a child passed the test, two parents added that it meant the child had been “taught to pass the test.” When a child failed, Josie commented it meant that not all the content had been taught before testing. Furthermore, failing did not mean the child would be less of a success in life. However, failing the test discouraged a student by focusing on his or her weaknesses and

not on a student's strengths. Sharon suggested that the discouragement, from failing, affected the success of a child by using her older children as examples. Her son was in the gifted and talented program in high school, but he had to take remedial math in college. If her son had been held back because of his weakness in math, the discouragement would have prevented him from earning the registered nurse position he holds today. The same example was used for Sharon's daughter who was also weak in mathematics, but she went on to graduate summa cum laude in psychology. Sharon expressed, "Just because of a weakness in one area, it doesn't mean students aren't going to be successful." Sharon favored the test for providing tutoring in the content area in which her child was weak. In fact, she appreciated the tutoring her daughter was receiving as a grade 8.5 student in mathematics, but she added, "don't hold them (eighth graders) back because of that (LEAP 21)." Sharon summarized her feeling of LEAP 21 as a single criterion for promotion with "my little girl always had a positive attitude until LEAP. When she didn't make it (the cutoff score for passing), it (failing the test) just crushed her."

Tom suggested grouping students based on their abilities. Two parents noted that students were all grouped together so that "better students will bring the others along." Tom disagreed with the practice of grouping all students together. He used the examples of "a bad apple in a bushel and the whole bushel is going to go bad" and "dumbing the curriculum down for groups of students, rather than, bringing everyone up."

Was the Test Fair? Parents explored the role of testing in the workplace and related LEAP 21 to it. Tom shared that testing takes place in apprenticeships to move a person to journeyman and for nurses to meet state board requirements. Drivers take tests to get a

license to drive a car. Tom shared that “a test is a test, and if you can’t pass a test, you don’t go on.” However, Sharon disagreed and expressed that driving was a privilege, and if you don’t pass the test, you are allowed to re-take the missed part. Sharon and Josie were pleased that students “were allowed to go on in school” as a grade 8.5 student. However, if children were prevented from going to the next grade as in grade retention, the discouragement affected their future success, and three parents disagreed with the practice.

Two parents expressed a need for parental support, and one parent added that inner motivation by the child was needed. Tom shared that he “came from a poor family,” and a person can still succeed with motivation. Therefore, the test did not impact rich and poor students differently.

Three parents said that if non-public school students received a high school diploma, then “why should they (non-public school students) not have to take the test?” And if the non-public school curriculum was so much better, students should not have any problem passing. For non-public school students to lack a testing requirement, the two parents suggested that non-public students were better than students in public schools. In effect, Tom and Sharon said that students were given a privilege, because they had the money to go to a non-public school. The non-public school students did not have to follow the same criteria as public school students. Three parents had not considered non-public schools for their children because of LEAP 21; however, they had considered home schooling, for as Josie said, “A host of other reasons.”

In all, City Street, White, grade 8.5 parents supported LEAP 21 to evaluate the improvement of Louisiana schools. They also supported LEAP 21 to tell what “kids

learn.” However, three of the four parents did not agree with use of the test as a single criterion of whether the student should graduate or be promoted. Discouragement from failing the test caused students to be less successful in seeking career opportunities. Two parents summarized with “don’t punish them (the children) because you want to know what they have learned.”

While describing perspectives by categories of parents is useful, an analysis across categories allows grouping of the data revealing relationships (Krueger, 1994, 2004) and showing the extent that opinions of parents differed by race, community type, and passing or failing the test. An analysis using themes takes place in the next section.

Analysis by Themes

The study used three themes in asking parents questions and in data analysis; reform intent, student learning, and test fairness. Discussions of the a priori themes and emergent themes are presented next.

Theme 1- Reform Intent

Focus group interviews were used to find the opinions of parents about why their children were administered LEAP 21. The analysis allowed comparisons of parent perspectives and revealed the extent of agreement or disagreement across categories of parents. Parent perspectives about the purposes of the test are presented in Table 20

Ten parents of students, who were both Black and White and passed and failed the test, said that LEAP 21 was used to find the academic level of a student and an additional six parents spoke of using the test to show the academic strengths and weaknesses of each child. A rural, Black parent of a ninth grader summed the feelings by saying, “I thought it (LEAP 21) was to see what they (the students) know and don’t know so they

Table 20

Test Purpose

(53) Parents Share Why Students Were Tested						
	(21) Rural		(9) Suburban		(23) Urban	
	(5) Pass	(7) Fail	(2) Pass	(3) Fail	(2) Pass	(10) Fail
Black	(1) The test	(2) The test	(2) The test	(2) The test	(2) The test	(1) The test
(29)	determined the	determined the	determined the	determined the	prepared students	determined the
	academic level of a	academic level of a	academic level	academic level of	for the future. (1)	academic level
	student. (3) The test	student. (1) The test	of a student.	a student. (1)	Parents did not	of a student. (6)
	identified students	prevented students		Parents did not	know.	Parents did not
	who passed and who	from graduating when		know.		know.
	failed. (2) Parents did	they could not read.				
	not know.	(3) Parents did not				
		know.				

(table con'd.)

	(5) Pass	(4) Fail	(1) Pass	(3) Fail	(7) Pass	(4) Fail
White (24)	(5) Parents did not know, even when probed for an answer.	(1) The test improved education. (3) The test identified student strengths and weaknesses.	(1) The test ended social promotion.	(3) The test identified student strengths and weaknesses.	(4) The test graded the teacher. (2) The test determined the academic level of a student. (1) The parent suspected a hidden reason.	(4) The test showed students could read. (2) The test was used to bring students to the national average.

Note. The numbers in parentheses refer to parents in each group and number of parents who held each opinion. Parents, who indicated no response either verbally or by gesture, cannot be assumed to agree or disagree with other participants.

can work on it (weak skills) to help them in subjects.” Two parents agreed that LEAP 21 was used to bring Louisiana education to the national average and testing was necessary to show that increases in student achievement were taking place. Nine parents shared that LEAP 21 stopped a student’s progress in school because of a lack of academic skills, measured through testing. The parent of a White, urban, grade 8.5 student shared the sentiment by saying LEAP 21 keeps “kids from graduating without the necessary skills.” However, the parent of a White, urban, ninth grader questioned comparisons with other states. The parent said the state had a unique culture and did not compare to other states. Two other parents in the same group suggested one of the ways the state differed from others was because of lower standards for learning. However, the standards were improving for “the failing (school) systems” in Louisiana. While 34 parent responses indicated a reason students were administered LEAP 21, 18 parents, across all categories except White suburban, responded to the question with silence. Parents then said, “I don’t know” or used the opportunity to add, “Why are we taking LEAP? That is a good question.” Interestingly, rural, White parents whose children passed the test offered no reasons as to why students were tested and expressed ““I have no idea”, “It replaced a test, didn’t it? “ and “there was the IOWA (test) and the CAT (test), don’t they still have those?” Because parents were asked this question at the beginning of the interview, an alternative explanation includes parents not initially feeling comfortable to share their understanding about LEAP 21 and the manner it contributes to improving the education in Louisiana schools. Five parents said that the test determined promotion. The parent of a suburban, White ninth grader shared the sentiment with:

“They (the state) instituted it at fourth, eighth grade, and high school so that kids wouldn’t be passed along without knowing anything. The test would say ‘yes’, they know what they’re supposed to know.”

The parent responses reflected a trust in the system, and testing was allowable if it benefited the child. In fact, no comments were made that parents did not want their children to be tested. Parent trust was strongest as indicated in the parent responses about the test being used to determine the academic level of the student and to identify a student’s strengths and weaknesses. Support of parent trust in the system also came from the few numbers of parents who suggested that testing was just learning if a student passed or failed the test.

Theme 2 - Student Learning

The second analysis allowed examinations of the extent students were making academic progress because of LEAP 21. The theme of student learning is used and the variation in the opinions of parents is presented in Table 21.

Eleven parents, eight of whom had White children, shared that LEAP 21 improved education as students were learning more. The parent of a White, urban, grade 8.5 student said, “I think it’s (LEAP 21 is) kind of forcing them (students) to do work that is ahead of them.” While the parent of a White, suburban, grade 8.5 student added that his child had to “overcome laziness in not completing problems, so the test improved his child’s learning habits.” Four of the parents spoke favorably about LEAP 21 tutoring for their child and viewed it as a means to improve education. The parent of a Black, rural, grade 8.5 student summed the sentiment saying, “They (the teachers) reinforce the remediation by tutoring in the afternoon. That helped my son a lot to retain the information.” However, 60 responses given by parents indicated that the effect of being administered a

Table 21

Educational Improvement

(53) Parent Perspectives on Whether the Test Improves Education						
	(21) Rural		(9) Suburban		(23) Urban	
	(5) Pass	(7) Fail	(2) Pass	(3) Fail	(2) Pass	(10) Fail
Black	(3) Yes, students	(2) No, the test	(2) No,	(2) No, students	(2) No,	(6) No, test content differed
(29)	learned. (2) No,	content differed	students felt	felt pressured	teachers did	from classroom learning;
	students were not	from classroom	pressured to	to pass the test.	not prepare	(4) teachers did not
	taught the content. (2)	learning. (1) No,	pass the test.	(1) No, test	the students	understand the content; and
	No, students felt	the students felt		preparation	for testing.	(5) parents did not
	pressure to pass the	stressed.		replaced other		understand the content. (9)
	test. (3) No, students			important		No, students “gave up” and
	did not learn because			learning.		thus did not learn. (6)
	of late test					Parents were frustrated and
	preparation.					(2) students were stressed.

(table con'd.)

	(5) Pass	(4) Fail	(1) Pass	(3) Fail	(7) Pass	(4) Fail
White (24)	(2) No, not all of the tested content was taught. (4) No, the pace of instruction was quick for learning.	(3) No, students felt pressured to pass the test. (4) No, not all students were capable of meeting the same expectations.	(1) Yes, students learned when instruction matched the test.	(1) Yes, students were expected to learn more. (2) No, students were stressed by the test. (1) No, test preparation replaced other important learning.	(4) Yes, students learned more to pass the test, but the students had to (7) “hurry” and (7) “worry.”	(2) Yes, students were forced to learn more. (1) No, test preparation replaced other important learning. (2) No, students were poor test takers and failing the test discouraged learning. (3) No, students were stressed by the test.

Note. The numbers in parentheses refer to parents in each group and number of parents who held each opinion. Parents, who indicated no response either verbally or by gesture, cannot be assumed to agree or disagree with other participants.

test was not improving the academic learning of a student for different reasons. Eighteen responses, most of which were made by parents of Black students in urban schools, suggested that students were not being taught the content which was tested, and therefore, LEAP 21 was not improving a child's education. The parent of a Black, rural, grade 8.5 student summarized with "the 'stuff' they practice isn't even similar to what was on the LEAP." The parent of a Black, urban, grade 8.5 student questioned the practice with "why give them (the students) something at the end of the year that they've never seen before all through the year?" Eleven responses of parents of White students indicated the pace or "hurry" of instruction was too fast for their children to learn the tested content. The parent of a White, rural, ninth grader shared parents' feelings and referred to the practice as "crunching it (the time)" by not spending the necessary instructional time on the content which was covered on the test. An additional 22 responses, across all categories of parents, cited the added stress or "worry about passing", which prevented their children from benefiting from improved instruction. The parent of a White, suburban, grade 8.5 student summarized the sentiment of the parents with, "Now it (tested content) is being pounded into them to get ready. I don't think the people who decided on making this test so high stakes realize the stress they are putting on children." Three responses reflected that test content replaced what they considered to be important content. A parent of a White, urban, grade 8.5 student explained with "I think the teachers are spending more time preparing the kids for the test than actually teaching them what they should be learning." Therefore, the parent felt the test was not improving the education of his child. The parent of a White, suburban, grade 8.5 parent said, "I just feel they (the students) learned more useful things before this LEAP became so

important.” Another three parents said the content was taught too late in the year to improve their children’s education, and two parents said that the test could not improve education, as some students were poor test takers. The parent of a White, urban, grade 8.5 student said, “Students know the material. They can apply it, but when it comes down to doing the standardized test, everyone is not a good test taker.” Four White, rural parents of grade 8.5 students blamed having students meet the same expectations as the reason that education was not improving. One parent shared that “you’re not going to find a kid that’s going to be able to do all of it and pass it with no problem. But they (the state) want all these kids to score the same things to pass; it’s not going to happen.” Another parent in the same group offered an explanation with “so you’re telling us that everybody is supposed to pass the LEAP, but not everybody is on the same educational level. You’ve got children that are learning disabled, dyslexic, that can’t read enough and they’re taking the same test.” Again, the parent responses indicated that the test was a useful tool for diagnosing academic problems, and parents supported testing when their children received help in the needed areas. Parents in all categories expressed that students were expected to learn more because of the test; however, factors of a mismatch between instruction and the content, the pressure on the students to perform on the test, and the requirement that all students meet the same expectations reduced the useful diagnostic purpose that parents gave for supporting the test.

Theme 3 - Test Fairness

Parent perspectives on the fairness of LEAP 21 were explored in several areas. The first examines the test disparities for groups of students; rich or poor, Black or White, and male or female. Next the opinions of parents are given about the lack of a test

requirement for non-public school students. Non-public schools in Louisiana, including private and parochial schools are not required to administer LEAP 21 to eighth graders for promotion to ninth grade, and parent perspectives were gathered on the lack of a testing requirement for students in the schools. The examination of LEAP 21 as a single criterion follows exploring parent perspectives on the high stakes consequences of LEAP 21, grade retention. Parent perspectives are then given for the ways timing of the test impacts students. Currently, LEAP 21 is administered to eighth graders in the spring of the year with students receiving the test results before summer. Parent perspectives were explored for the time in the school year that testing takes place.

Parent Perspectives on the Disparate Impact. Parents gave their opinions on the manner LEAP 21 affected students who came from affluent families versus those who did not. Disparities were also explored for students by race and gender. The differences testing causes on these groups of students are given in Table 22.

Sixteen parents of Black students expressed that the wealth of a family made a difference. The parent of a Black, rural, grade 8.5 student said “rich students were exposed to more,” and students in wealthy families would know more because “they have the resources.” The resources included being able to provide help with homework. The parent of a Black, urban, grade 8.5 student described what trying to help your child was like when a parent lacked finances and academic resources. “I’m their tutor; I’m the best they got” and “How am I going to help my child with homework, if I don’t understand what they’re doing?” were the comments made. Another parent in the same group added that students in wealthy families are able to travel more. The parent of a Black, urban,

ninth grade student added, “Money makes the world go around” when resources were needed

On the other hand, 16 parents of White students said there was no difference by family affluence. Six parents of White students said that the level of home support made a difference as a parent of a White, suburban, grade 8.5 student expressed, “You’re either going to have a poor mother who’s going to sit there and go page by page or you’re going to have a rich mother do the same. So it (being rich) has nothing to do with it (the test).”

Eleven parents of Black students said testing impacted Black students more than White students. A parent of a Black, urban ninth grader summarized the sentiment with “Look at the numbers. I bet it’s ours (Black students) over theirs (White students). We’re up (on a scale); they’re down.” Three parents of Black children in rural schools who passed the test found less of an impact on Black students. “I don’t see any difference. I figure the percentage is about the same with passing and failing. It has nothing to do with race.” Two additional parents of Black students in suburban schools shared the sentiment. An alternative explanation, for not finding race to be a factor by the five parents, was the parents did not feel comfortable expressing race differences when the interviewer was not the same race as them. Two parents of White, rural, grade 8.5 students recognized the manner testing differed by race of the student by suggesting LEAP 21 was a “White man’s test.” The parent blamed the spoken language of Black students, who were forced into housing projects because of finances, as the reason which prevented students from “doing so well” on the test. Nine additional parents of White students, who both passed and failed the test expressed that the test did not impact Black and White students differently. The parents offered no reasoning to support their claim. Four parents talked

Table 22

Test Disparities for Groups of Students

(53) Parent Perspectives on Disparate Impact – Rich/Poor, Black/White, Male/Female						
(21) Rural		(9) Suburban		(23) Urban		
	(5) Pass	(7) Fail	(2) Pass	(3) Fail	(2) Pass	(10) Fail
Black	(2) Students in wealthy families had an advantage. (3) Race did not make a difference. (2) Girls were smarter.	(3) Students in wealthy families and (2) White students were advantaged (3) Boys and girls were affected the same.	(1) Wealthy parents afforded summer school easier. (1) Student differences of race and gender did not matter.	(1) Wealthy parents afforded summer school easier. (1) Student differences of race and gender did not matter.	(2) Students in wealthy families had an advantage because the students traveled more. (2) Black students failed at higher rates.	(7) Students, in wealthy families and who were White, had an advantage. (5) Boys and girls were affected the same by the test.

(table con'd.)

	(5) Pass	(4) Fail	(1) Pass	(3) Fail	(7) Pass	(4) Fail
White (24)	(3) Home support made a difference and not being rich/poor or Black/White. (2) Girls did better on the test. (1) A child's personality determined test differences.	(4) Differences did not exist for students in rich or poor families and for (3) boys or girls. (2) The test was a "White man's test."	(1) Boys were lazier in test performance.	(3) Differences did not exist between students in rich/poor or Black/White families. (2) Race did not make a difference. (1) Boys were lazier in test performance.	(3) Being rich/poor or Black/White made no difference. (1) Girls cried when they failed.	(3) Inner motivation and home support were important and not being rich or poor. (3) Teaching of the content on the test needed to be adjusted to assist academically poor students.

Note. The numbers in parentheses refer to parents in each group and number of parents who held each opinion. Parents, who indicated no response either verbally or by gesture, cannot be assumed to agree or disagree with other participants.

about girls performing better on the test than boys, because as one parent of a White, rural ninth grader suggested, “I think girls feel more pressure for some reason than boys do.” The boys are just a little more relaxed about taking the test.” Therefore, girls did better, because they focused more on the test. Two parents blamed boys for being lazy as a reason that boys did not perform as well as girls on the test. Fourteen parents said that boys and girls were affected the same by the test. A parent of a White, rural, grade 8.5 student summarized the sentiment with “I don’t think it makes a difference whether it’s a boy or girl. I think it depends on that individual child.” Overall, parent perspectives on the disparate impact of LEAP 21 differed by race of the child more than community type or whether a parent’s child passed or failed the test.

Testing for Public/Non-public School Students. Parents shared their opinions about the lack of a testing requirement for non-public school students. Differences by categories of parents are given in Table 23

. Eighteen parents of Black students, located across all three community types, said the practice of not requiring non-public school students to test was an unfair practice.

Whereas, seven parents of White students said the practice was unfair. A White, rural, ninth grade parent explained that non-public school students “should take the LEAP test just like our ‘regular’ kids do. They’re (the non-public school students) getting a state diploma.” The parent of a Black, suburban, ninth grader also acknowledged, “They (non-public schools) are getting funds from them (the government) and the students should be able to take the test.” Seven parents of White students admitted that they would use the practice of enrolling in a non-public school to avoid LEAP 21 for their children with the

Table 23

Disparities of Public and Non-Public School Test Requirements

(53) Parent Perspectives on the Lack of a Test Requirement for Non-public School Students						
	(21) Rural		(9) Suburban		(23) Urban	
	(5) Pass	(7) Fail	(2) Pass	(3) Fail	(2) Pass	(10) Fail
Black	(3) Yes, it is an unfair practice.	(4) Yes, it is an unfair practice. (3) No, parents did not consider the practice.	(2) Yes, it is an unfair practice	(1) Yes, it is an unfair practice.	(2) Yes, it is an unfair practice	(6) Yes, it is an unfair practice.
(29)	(1) A parent was not aware of the practice.	(2) Yes, parents would use the practice.	(2) No, parents did not consider the schools.	(2) No, parents did not consider the schools. (1) Yes, a parent would use the schools.	(2) No, parents did not consider the schools.	(2) A parent was not aware of the practice. (3) Yes, parents considered the schools.
	(1) Yes, a parent would use the practice.					

(table con'd.)

	(5) Pass	(4) Fail	(1) Pass	(3) Fail	(7) Pass	(4) Fail
White (24)	(1) Yes, it is an unfair practice. (1) A parent was not aware of the practice. (5) No, parents would not use the practice.	(3) No opinion was given, but parents spoke of students that they knew in non-public schools. (1) Yes, a parent would use the practice. (1) No, a parent would not use the practice.	(1) Yes, it is an unfair practice. (1) Yes, a parent would use the practice.	(2) No, parents would not use the practice. (2) Non-public school students were required to test to enter public schools.	(2) Yes, it is an unfair practice. (2) Yes, parents would use the practice. (3) No, parents would not use the practice	(3) Yes, it is an unfair practice. (3) Yes, parents would use the practice.

Note. The numbers in parentheses refer to parents in each group and number of parents who held each opinion. Parents, who indicated no response either verbally or by gesture, cannot be assumed to agree or disagree with other participants.

parent of a White, suburban, ninth grade student summarizing the sentiments with “if my kid would have failed in March and failed it (LEAP 21) in July, my kid would be in private school.” An equivalent number of parents of Black students said they would use non-public schools to avoid LEAP 21 requirements for their child. The parent of a Black, rural, ninth grader summed the sentiment with “I would rather see my child go (to a non-public school) and know she’s going to get a high school diploma than not get a high school diploma because she failed the LEAP test.” The parent of a Black, urban, ninth grader talked about what having money for tuition would do for her child:

If I had the money...I would send my son because I know what a struggle it is for a Black male. I would send him for the opportunity not to have to go through low self-esteem as far as being told you’re going to fail because you flunked the LEAP test.

Thus, 14 parents of students, both Black and White, in the three community types considered using the practice of enrolling in a non-public school to avoid LEAP 21 consequences for their child. The parent of a Black, urban, grade 8.5 student said, “Yes, I did. I considered it (sending the child to a non-public school), but I can’t afford it.” Thus, parents of students in the three community types, while considering non-public schools for their children, also recognized that family affluence affected their decision.

Suggestions are made that parents who could afford to send their child to a non-public school to avoid LEAP 21 consequences were using the practice, and that finances also affected parents who responded negatively to considering the schools for their child. The extent, that finances affected parents who responded positively or negatively, was not determined in the present study.

Single Criterion. Parent perspectives were also analyzed for use of LEAP 21 as a single criterion for promotion. The feelings of parents were similar across groups, except for a few opinions.

Three parents said they agreed with use of LEAP 21 as a single criterion for promotion. The parents who agreed were White: two whose children attended a school in a suburban community and the other whose child attended a school in an urban community. Two parents supported the test for advancement through the workforce and through the military ranks. The third parent supported the test when multiple opportunities were provided to pass LEAP 21, and acknowledged her access to child specific tutoring. Forty-five parents responded negatively to the question of using LEAP 21 as a single criterion for promotion. The parents expressed a depth of feeling through frequency and intensity of responses about the manner the test was used with their children. A Black, rural, ninth grade parent summarized the sentiment with “I don’t think it is fair at all. They (the students) go to school all that time, and then, it depends on one test whether they get promoted or not.” Likewise, a White, rural, ninth grade parent described, “It’s a horrible situation to say that I have to pass this test no matter what my record, my behavior, my discipline skills.” The parent further described the use of the test as a single criterion on students as “even if they (the students) struggle all year long and make good grades, it’s all for naught. They have to pass this one miserable test.” The parent of an urban, Black, grade 8.5 student expressed her own confusion when a student was passing all year. “You (the teacher) told me he was learning. You told me he was making these grades. Then what happened with the test? That’s going to make that child not want to go back to school.” The parent felt misled when LEAP 21 results did not

match report card grades. The parent of a rural, Black, ninth grader spoke about the way school accountability was tied to student performance on a test by saying:

If they want to bring the schools up to the national average and they want to be able to evaluate what the kids have learned, fine, then do that. Then give the LEAP test. But it should not be a measure of whether they (the students) graduate or whether they are promoted.

Two parents of Black, urban ninth graders added that the consequence of high stakes testing was “one more factor of losing the child to the street.” The parent of a White, rural, grade 8.5 student summarized the four parents in her focus group saying “this is just a test for one week. It (passing) should go by what they (the students) do all year long.”

The only White parent of a ninth grader in the suburban community to participate in the study indicated why parents disagreed with the test as a single criterion. The parent metaphorically (Krueger, 1994) described the effect of failing the test on students as being “hit over the head with a hammer.” The parent took precautions to guard her child against the discouragement of failure, by registering the child at a non-public school as a buffer against the discouragement.

Timing of LEAP 21. The test is administered to eighth graders in the spring of the school year. Students receive results in early May, and then failing students are given the option of summer school attendance and the opportunity to re-test in July. Parent perspectives are given on LEAP 21 impacts on their children as the students enter high school and as they complete their eighth grade year. The perspectives are summarized in Table 24.

Two parents of Black urban, ninth graders stated that the time of the year that the test administration took place did not matter because as one parent expressed, “I believe the

Table 24

When LEAP 21 Testing Occurs

(53) Parent Perspectives About When Students Take LEAP 21						
	(21) Rural		(9) Suburban		(23) Urban	
	(5) Pass	(7) Fail	(2) Pass	(3) Fail	(2) Pass	(10) Fail
Black (29)	(3) When the students failed the test, they opted for a GED.	(2) The test administration took place late in the year. (1) Students in 10 th grade were given multiple chances to pass the exit exam, but not eighth graders who took LEAP 21.	(1) Administer the test at the beginning of the year, so students can work on their weaknesses.	(2) The test needed to be given at the beginning of the year, so students could work on their weaknesses.	(2) Students would pass or fail regardless of the time of year. (2) Students quit school because of LEAP 21 timing.	(6) Students quit school because of LEAP 21 timing. (2) Students opted for a GED.

(table con'd.)

	(5) Pass	(4) Fail	(1) Pass	(3) Fail	(7) Pass	(4) Fail
White (24)	(4) The test administration took place close to eighth grade graduations.	(2) When the students failed the test, they opted for a GED. (1) Students quit school because of the timing of LEAP 21.	(1) Eighth graders needed multiple opportunities to test as 10 th graders were given on the high school exit exam.	(3) Parents did not comment on when the test was administered.	(3) Students learned of the pass/fail consequences late in the school year.	(1) Students learned of the pass/fail consequences late in the school year. (1) Students quit school because of the timing of LEAP 21.

Note. The numbers in parentheses refer to parents in each group and number of parents who held each opinion. Parents, who indicated no response either verbally or by gesture, cannot be assumed to agree or disagree with other participants.

results would be the same.” In effect, students would pass or fail no matter what time the high stakes test was given. Thirteen parents admitted that the administration of the test was “late” in the school year. Four parents of White, rural, ninth graders said that the test administration was late for determining whether their child would graduate from eighth grade. One of the parents summed the feeling of parents in her group with “that’s what we were going through at that time period. They (the children) worried about their graduation. Would they be with their friends? Am I going to graduate? They were on pins and needles.” No parents in suburban or urban schools spoke about eighth grade graduations. The suggestion is that the practice of graduating to high school did not take place in the larger school systems. The timing was said to be late for finding out about whether a student was promoted to the ninth grade. The parent of a White, urban, grade 8.5 student summarized the feelings with “it was a lot of stress, because we didn’t know until at the very end of school. We wondered if we were going to get promoted or were we going to be held back?” Three parents of White, urban ninth graders and two parents of Black, rural, grade 8.5 students also expressed that the test administration was late for knowing that their child failed. Three parents of children who passed and failed the test in suburban schools agreed on why the test was considered late. Students did not have time to work on their weaknesses when results were given as the school year ended. One parent in the group summarized the feeling by saying, “Take the test at the beginning of the school year. Then you know where (academic level) they (the students) need to be. If you just give it to them at the end of the year, that doesn’t mean anything.” The parent of a Black, rural, grade 8.5 student also spoke about how the timing of the test was late. She offered that administering the summer re-test left students confused as to their grade

placement for the upcoming school year. She said, “When they (school administrators) end that summer school, we have to wait. He (the student) is standing there. Where do I go? I could not put him in ninth grade. He had to go back in the eighth grade.” Thus, the test results from the summer re-tests were not available until the schools began opening in the fall.

Seventeen parents spoke about students dropping out of school or getting a Graduate Equivalency Diploma (GED) because of being administered LEAP 21, with its high stakes consequences. Eight parents of Black students in urban schools said students dropped out of school when they failed the test. One parent in the group expressed the sentiment with “you’re going to have more and more children dropping out. It’s sad to say, if you look at it, it’s mostly the Black kids.” Five parents of Black students, three in rural schools and two in urban schools, said students considered getting a GED because of grade 8 LEAP 21 testing. The parent of a Black, rural, ninth grader summed the sentiment with “Why should I send my child to school forever? She gets close to the door and passes all the classes. She has to get a GED because she can’t pass the stupid LEAP test that the state decides she needs to have.”

The number of parents of White students who considered lowering their expectations from a high school diploma to getting a GED or quitting schools were fewer than parents of Black students. Two parents of White, rural, grade 8.5 students spoke of students getting a GED because of the timing of test administration. One of the parents in the group described the attractiveness of the GED option.

If it becomes too much of a hassle, I’m going to switch him to the GED program where you don’t have to worry about the LEAP test period. He can still get a high school diploma, graduate on stage, do everything like the other kids do but not have to take these standardized tests.

Also, two parents of White students who failed the test, one in a rural school and one in an urban school, spoke of students quitting school. The parent of a White, rural grade 8.5 student spoke about the age of 16 and the reality of students quitting school as they approached the age. “So what is her life going to be like when she doesn’t finish school? I think she is going to quit when she turns 16.” Parents of students in suburban schools did not speak of their child getting a GED or their child dropping out of school. Here, the suggestion is made, that affluence and access to resources for private tutoring or enrollment in a non-public school, offered options to parents of students in suburban communities that less affluent parents lacked.

The parent of a White, rural, ninth grader offered a reason as to why timing of the test was especially hard for eighth graders. Tests were administered at a time when students were going through “puberty and social changes” and at a “time that a parent begins to lose control of the child because of peer pressure.” The parent of a Black, urban, ninth grader elaborated on parents losing control of the eighth grader:

After a certain age, you lose your hold on your child. Yes, you do because of peer pressure. In society today, the Black male is easily influenced to the streets as far as drugs, whether it’ using or selling. I feel that (failing the test) could cause a lot of our kids to quit school and be subject to that environment.

Hence, parents, even though well intentioned, had less influence on the choices their children made during adolescence. This finding suggests that timing of the test becomes more important in reducing the disparate impacts, especially for Black students.

Summary

Parents had varying levels of understanding about why students were tested. The understanding ranged from none to parents saying the test was used to show the academic

performance of a child. Parents also offered that testing took place to end social promotion.

Few parents spoke about students who learned more because of LEAP 21. Parents, who said the test was used to improve education, did so because more learning was expected of students. A large percentage of parents did not think LEAP 21 improved education because the students were not taught the test content, the pace of instructional delivery did not allow for student differences, the stress experienced by students interfered with learning, and the preparation for the test replaced important content.

Black parents said that family affluence made a difference in students taking LEAP 21, allowing them access to more resources and opportunities. Parents of White students said that family affluence made no difference for the way students performed on LEAP 21, but rather, family support did. Fewer parents spoke about the test impacting Black and White students differently. Eleven parents of Black students and two parents of White students said more Black students failed the test than White students. Parents said boys and girls were affected similarly by LEAP 21 and described each child as unique to the manner that the test affected them.

Parents of Black students found the lack of a test requirement for non-public school students was unfair at rates higher than parents of White students. Also, parents of Black students who disagreed with the practice were pre-dominantly located in urban schools. Parents of Black students considered sending their children to a non-public school to avoid LEAP 21 requirements at a rate equivalent to parents of White students. Parents in all three-community types recognized that affluence affected their decision about use of a non-public school for their children. Also, suggestions were made that parents who could afford to send their child to a non-public school were doing so and some parents who did

not consider the practice lacked the finances. The extent of these considerations was not determined in the present study.

Support for use of the test as a single criterion came from three parents of White students who favored testing for advancement or if multiple opportunities were given to pass. The remaining parents disagreed with use of LEAP 21 as a single criterion for promotion. A suggestion was made that parents disagreed with use of the test as a single criterion because students became discouraged when they failed; and, the discouragement created a barrier that most poor and minority students lacked support to overcome. Hence, students quit school or opted for a GED when they failed. Support, that a parent previously gave for LEAP 21 to identify student strengths and weaknesses and to determine the academic level of a student, was withdrawn when the test was used as a single criterion. Also, trust, which a parent had in the school system, was diminished when the child received passing grades during the school year, but failed eighth grade, because the test was used as a single criterion.

As to administering the test in the spring of the school year, a limited number of parents of Black students found timing of the test was not a problem. However, parents said the test administration occurred late in the school year for determining attendance at eighth grade graduation ceremonies and attendance at a particular school campus after the summer re-test. About a third of the parents of students in rural and urban school communities spoke about students dropping out of school or getting a Graduate Equivalency Diploma (GED). Parents of students in suburban communities did not speak of the two alternatives because of failing LEAP 21. The finding suggests that parents of mostly White students in more affluent suburban community types did not consider less than a high school diploma for their children.

CHAPTER 6. DISCUSSION, CONTRIBUTIONS, AND RECOMMENDATIONS

The current study addressed three purposes. The first was to present the evolution of standards based reform (SBR) as a background for why testing became high stakes, especially in Louisiana. The second purpose was to describe test score differences for Black and White eighth graders on *LEAP for the 21st Century* (LEAP 21). Use of community types facilitated elaboration of test score differences across rural, suburban, and urban school locations. The third purpose was to provide further insights into the effects of testing on students in the perspectives of parents. Chaos theory was used as a lens to examine and understand how the multi-leveled, complex, and disorderly substance of the phenomenon, high stakes testing of SBR, impacted students. The lens allowed examination of SBR as a part of a recursive-dissipative cycle revealing “equality” returning as a focus previously held by “excellence” and “accountability” (Gordon & Bonilla-Bowman, 1994). The present study supported an equality or fairness focus in examining high stakes testing impacts through disaggregating the data by race and community type and by including questions on the fairness of the test during the focus group interviews.

Discussion

Standards based reform (SBR) was designed to set high academic standards for students in Louisiana and to end the practice of social promotion. State officials utilize LEAP 21, a high stakes standardized test to measure progress in the reform effort. The quantitative analysis of test score differences between Black and White students in the current study supported prior research revealing Black students do not test as well as White students (Gladfelter, 2000; Lewis, 2000) and that test scores for promotion have a

disparate impact on racial and ethnic minority students (Hauser et al. 2000; Johnson & Johnson, 2002). Also, the study contributed to the literature by supporting previous research by Johnson and Johnson (2002) in disagreeing with use of the test as a single criterion for student promotion. While Johnson and Johnson (2002) used Louisiana fourth graders and the method of case study, the current study examined eighth graders and used an analysis of test score differences and parent focus group interviews.

Statistically significant test score differences, often attributed between students in urban schools, were found in all school community types including suburban schools where resources are considered to be more accessible. Positive test score differences, favoring White students, were also found across all achievement levels. Interestingly, parents, through the focus group interviews, did not express the extent of the differences as revealed in the quantitative analysis. The descriptions were limited to recognizing that Black students failed at higher rates, but not that Black students were underrepresented in higher achievement levels. Also, parents suggested how the test score differences impacted students. Thus, the use of the two phases was important in the present study.

SBR implementation becomes having Black students perform better on standardized tests as well as improving education through state content standards. The effect is a disparate impact on Black students, by using the tests as the means to measure improvement in the reform. The multileveled description of SBR implementation allowed some understanding of why the practice takes place. Community stakeholders are also involved and K-12 schools are subject to tests to satisfy stakeholders, whereas, colleges are not. This suggestion supports research by Rose and Gallup (2004) of the public's support of high stakes testing. Consequently, Black students are disparately impacted

when SBR progress is assessed through a single criterion standardized test to satisfy community stakeholders, while findings from this study suggest parents, of both Black and White students, disagreed with use of the test in this manner.

Black students failed LEAP 21 at rates higher than White students. Parent perspectives revealed the discouragement that students felt when failing the test. The disproportionate failure of Black students caused the students to consider a Graduate Equivalency Diploma (GED) or quit school at rates higher than White students. The findings add to previous research by Reardon (1996) of the increased feelings of dropping out in use of eighth grade tests. Because SBR uses LEAP 21 to ensure high academic standards are met, the reform has been ineffective for Black students, because students are not in school to meet the standards. This finding was supported in the parent interviews of Black students affecting students in rural and urban schools more than students in suburban schools.

Smith et al (2004) use the term “political spectacle” for the state of political life in an era involving economics, communication, and language. Here policy instruments of high stakes tests become a cheaper and quicker reform alternative that also provides bureaucratic control over teaching and learning (Smith et al., 2004). High stakes tests create a “pseudo-reform,” according to Smith and colleagues through communication and language of politicians. The effort is a pseudo-reform because it lacks other necessary elements such as the professional development of teachers, increasing teacher pay, reducing class sizes, creating small learning communities, and above all, making sure that students are provided with an opportunity to learn. Poorer states may use high stakes tests for credibility in the reform process, but, in actuality, few meaningful reforms exist.

The passage rate improvement for Black students in magnet schools suggests higher academic expectations and quality teachers, often attributed to the schools, have a positive impact on the academic achievement of Black students in urban schools. While testing differences existed across all community types, providing high expectations, quality teachers, and resources (Goldring & Smrekar, 2002) to reduce the achievement differences between Black and White students is supported by the present study.

Parents wanted their children to learn and testing was a part of the process with parents recognizing LEAP 21 was a means to determine a child's academic level and a child's strengths and weaknesses. Parents suggested that testing take place earlier in the school year allowing students to work on their weaknesses. The support given by parents for the test as a useful tool in diagnosing academic levels of a student suggests a trust that parents held about the educational system. The lens of chaos theory, to view the educational system from a distance, reveals the decision to pass or retain a student is made at the state level and not by the local classroom teacher. The present study suggests use of a test with an arbitrarily set cutoff score for promotion negates the academic performance of students in a local system, especially when students had passing grades all year. Therefore, the trust, parents held for the educational system, was diminished because of the conflicting reports on student progress that parents receive.

The present study examined differences by gender through use of a question in the focus group sessions. Data were limited by use of one question and the results were limited by the lack of data to fully explore the impact on males and females.

Parents tied affluence and non-public school attendance together by speaking about being able to afford tuition and in a few cases speaking about both topics when one topic

was mentioned. Parents considered lack of a test requirement for non-public school students as unfair. Because non-public schools are attended by pre-dominantly White students, the study suggests that meeting a single criterion requirement that more affluent, White students do not have to meet disparately impacts Black students.

Parents explained that the tested content was not always taught in preparing students for LEAP 21. The Louisiana Department of Education (LDE) in 2004 released grade level expectations (GLE), providing specificity to content standards, and a model curriculum framework, providing activities to be taught to cover the GLEs. Still missing is the assurance that the content is taught, or simply stated, that students are given the opportunity to learn. Therefore, the current study suggests a need for monitoring the delivery of standards based instruction to facilitate students learning the tested content

In the current study, parents told about their lack of knowledge in how to explain the tested content of LEAP 21 to their children. The mathematics content differed from what parents had learned in school as revealed in the focus group interviews. Thus, parents were unable to help their children when given tutoring materials. Therefore, the study supports the selection of mathematics content of which parents were unfamiliar added to the disparate failure rate. Consequently, the study contributes to the literature in supporting findings that the standards and constructs state officials select matter (Linn, 2000; Willingham & Cole, 1997). Previous research by Linn (2000) included content areas of reading, history, and geography and found that boys or girls were advantaged over the other in each content area. The present study suggests that a group of students, mostly affluent White students, is advantaged by state officials selecting mathematics content which differs from the mathematics parents learned in school. The advantage

occurs because more affluent parents are able to get their children help with mathematics; whereas, less affluent parents lack the resources to do so. The study suggests this finding even when the local schools provide LEAP 21 tutoring sessions. Thus, the selection of mathematics content of which parents are unfamiliar, disparately impacts poor, minority students.

Policy Considerations and Recommendations

The first policy consideration provides monitoring the instructional delivery in classrooms for the effectiveness of standards based teaching and learning. An alternative to the use of a standardized test is to follow the progress of SBR through setting specific goals for learning and meaningful school improvement plans (Commission on Secondary and Middle Schools, 2004). The Commission on Secondary and Middle Schools (2004) sets standards for school organization and instruction. The standards are used in an accreditation process through an on-going peer review process, and they provide the process of monitoring instruction in schools to assist in the process of ensuring that students are provided an opportunity to learn. Raising the academic achievement of students in Louisiana classrooms needs to occur by providing more than the manner the courts have defined an opportunity to learn, teachers recognizing that tested skills were ones they should teach (*Debra P. v. Turlington*, 1981; Heubert & Hauser, 1999). In support of the district assistance team (DAT) and the school improvement team (SIT) used by Louisiana schools to improve classroom instruction and raise student test scores, the peer review team uses educators from outside the district to commend school practices and to make school improvement recommendations. This policy consideration, of an external review team monitoring for instructional delivery, is made based on parent

perspectives that instruction did not match the tested content, and it provides a similar process for K-12 schools that colleges use for monitoring educational goals. The monitoring follows a format similar to the DAT model, but includes an external review process to monitor standards based instruction and school reform efforts. Thus, the public trust that postsecondary schools enjoy can also exist for K-12 schools, without the use of a high stakes promotional test.

The second policy consideration is the use of multiple criteria for student promotion because of the disparate impact of the high stakes consequences of LEAP 21 revealed in the first phase of the present study. The policy consideration supports the test to identify a student's strengths and weaknesses in meeting high academic standards, but the policy removes the test as a single criterion for promotion. The support, for this policy came from parent interviews. The support was strongest when students had passing grades throughout the year. Therefore, the test is used as a tool to diagnose and direct instruction as needed by the students. The policy restores the trust of parents in the school system by the state recognizing grades of the local system as another criterion needed for promotion. Hence, the state supports the decisions made by the local system in the promotion of students, and the local school system implements high academic state standards using the state test as one criterion in student promotion and as a measure for meeting the standards.

When multiple measures are considered in high stakes decisions affecting public school students, similar high stakes measures, including standardized tests are warranted for non-public schools. One group of non-public school educators, the National Catholic Educational Association (NCEA), offers a statement on accountability and assessment in

Catholic education (NCES, 2004). The statement allows standardized tests as one measure of being accountable to the Catholic community and funders. As recipients of tax dollars for textbooks and transportation, Catholic schools in Louisiana are partially funded by tax payers. However, students graduating from Catholic high schools are not required to pass the state exit test, though they, nevertheless receive a diploma carrying all the rights and privileges as a state diploma. Historically in Louisiana, proposed legislation to dismantle the dual system of schools is countered by strong non-public school lobbying efforts. These lobbying efforts have, to date, prevented dissolution of the dual educational system operating in the state. Consequently, a third policy consideration derived from this study recognizes the tax payer as a funder of Catholic schools and therefore, according to the NCES position, entitled to accountability information through the use of standardized tests of student achievement. Recognizing that the current dual education system advantages more affluent, White students adds to the need for using accountability measures in both systems. Inclusion of multiple measures for grade promotion and high school graduation levels the playing field notwithstanding the dual education system.

The fourth policy consideration is to work with parents. They need reminding of the reasons their children are administered LEAP 21. While the information may have been given previously, parents of students taking the test focus on passing, and are therefore, unable to see how the test is part of improving the education of Louisiana students. Consequently, parents need the information to support their children through testing.

Because the test affected the time students had to work on academic weaknesses, the attendance of students at eighth grade graduation ceremonies, and the promotion of

students to the high school campus, parents encouraged offering more chances for students to pass. The test was compared to the high school exit test in which students begin testing in their 10th grade year to graduate in their 12th grade year. The eighth-grade test is given in the spring of the school year with one additional chance to test in July before being retained or assigned to grade 8.5. Timing the test earlier, possibly beginning in the seventh grade year, would allow instructional help for students without interfering with eighth grade graduation ceremonies and moving to a high school campus. Administering the test, whereby students have multiple opportunities to pass, is a policy consideration supported by the present study.

The sixth policy consideration is reviewing the quantitative data and the impact that testing has on Black students. Currently, students need to score Approaching Basic to pass ELA and mathematics. An examination of the data from this study revealed how the largest percentage of Black students was found in the failing category, Unsatisfactory, for mathematics and in the Approaching Basic category for ELA. Should the cutoff score be raised to a minimum score found in the Basic achievement level, the largest percentages of Black students would fail ELA as well as mathematics. However, the largest percentages of White students would pass the two sub-tests. Therefore, raising the cutoff score should not be considered until support is provided to move larger percentages of Black students into the Basic category.

Five recommendations flow from the present study. The first is to replicate the quantitative analysis using school community types identified from the 2000 census. This study used the 1990 census data for Louisiana school community types, the only available data for the 2003 LEAP 21 administration. However, the LDE was upgrading to

the 2000 census based school community types, and a replication of the quantitative analysis using the upgraded identification is recommended.

A second recommendation includes identifying schools that exceed the average performance for Black students across the three community types for follow-up case studies. Follow-up research would allow best practices of instructional delivery to be identified for improving test score performance for Black students. Because Black students were underrepresented in the higher achievement levels, the identification of best practices is needed in schools whereby greater than average percentages of Black students scored in the higher achievement levels. In addition, while research supports teachers changing their instructional delivery to best meet the needs of testing (Dorn, 1998; Madaus, 1988, 1991; Shepard, 1991; Smith, 1991; Smith & Rottenberg, 1991), observing and interviewing teachers regarding the instructional practices they use in schools where Black students meet high achievement levels in Louisiana is also needed.

While school districts vary in the way they organize to facilitate reforms, a third recommendation would be to identify district-level practices that best support state standards based instruction in classrooms. The sample, for the suburban focus group in the present study, reflected a tightly coupled system by the way the district controlled access to parents. Using the lens of chaos theory, to view SBR from a distance, allows tracing the multileveled implementation from state design to local delivery. The study suggests that district organization facilitates or hinders the communication of standards to teachers and instruction to learners. Examining Louisiana school district organization to facilitate successful implementation of state led reforms would be useful across the three community types. Parents of urban students believed that students lacked instruction

specific to the standards, and therefore, to passing the test. Hence, identifying district practices of organizing to support clear communication and delivery of state standards is needed.

The present study lacked methods to fully examine test score differences by gender. The use of one question in the focus group interviews did not allow full examination of the manner boys and girls are affected by standardized testing. The study lacked a quantitative analysis of test score differences by gender and sampling to accommodate gathering parent perspectives homogeneous by gender of their child. Therefore, a fourth recommendation for further study is to examine the differences by these methods.

A fifth recommendation for research is to examine student perspectives of high stakes testing. Student perspectives are called one of the most important variables in educational research by Caporrimo (2001) and the missing voice in educational research by Cook-Sather (2002). Even though there has been a call to include student perspectives (Erickson & Schultz, 1992; Phelan, Davidson & Cao, 1992), the response has been sparse (Cook-Sather, 2002). Therefore, interviewing students about high stakes testing in Louisiana is a recommendation.

Finally, more study is needed in identifying practices whereby passage rates of Black students improved on LEAP 21 at rates greater than White students when magnet schools were included as opposed to when they were not. Magnet schools may set higher expectations for students and Black students may respond by scoring higher. As an alternative explanation, because magnet schools organize around a content area, Black students' achievement increases as a response to student interest and engaging content.

Therefore, a sixth recommendation includes exploring the role of magnet schools in improving the scores of Black students on LEAP 21.

Summary

In this chapter, findings were discussed showing the use of a standardized test, LEAP 21, to measure the progress of SBR implementation disparately impacts Black students who do not test as well as White students. The test is used as a single criterion for promotion for eighth graders, and thus, the test subjects Black students to disparate rates of failing while showing the progress of implementing the standards.

Black students who failed the test considered quitting school or getting a GED because of the discouragement caused by the failure. Black students benefited from attending a magnet school and the current study suggested that high expectations and quality teachers reduce achievement differences between Black and White students.

Parents supported testing to reveal student academic strengths and weaknesses and believed in the educational system to help students in the weak academic areas. However, parents withdrew their trust when students received passing report card grades, but they failed the test.

Parents felt not requiring non-public school students to pass LEAP 21 was unfair, and the parents expressed that family affluence and being a White student contributed to the lack of a testing requirement. Parents spoke about students not receiving the appropriate instruction to be successful on the test while the state is providing more specificity about what content is covered and providing a curriculum for a teacher to teach. However, the provisions lack assurances that instruction and learning needed to pass the test, occurs because of the added initiatives.

Policy considerations were made including monitoring instruction for meeting state standards through recommendations of the Commission on Secondary and Middle Schools (2004) peer review process. Using multiple criteria for promotion is another policy consideration as well as providing a similar design of the promotional criteria for students in non-public schools. Also, parents need to be reminded of the reasons their children are administered LEAP 21, to maximize parental support in test preparation.

Testing needs to take place at a time whereby students have multiple opportunities to pass and students and parents receive test results timely. Also, raising the bar for eighth graders by setting a higher cutoff score cannot take place until instruction necessary to pass the test and timing the test administration allows the largest percentage of Black students to meet a passing score, as it allows for White students.

Recommendations for further research include a replication study categorizing school community types of students by the 2000 census data. Case studies of schools that exceed average test performance of Black students is also needed to identify instructional practices allowing attainment of higher levels of achievement on LEAP 21 by the students. More research is needed to identify optimum district organization to facilitate state led reforms. Research is needed to fully explore differences by gender in high stakes assessment. A quantitative analysis and sampling for parent interviews, homogeneous by gender of the child, is needed to fully describe testing differences for boys and girls. Allowing students into the reform conversation is also a recommendation by gathering their perspectives on assessments for academic achievement. Finally, exploring the role of magnet schools in improving Black student performance on LEAP 21 is a recommendation of the current study. This summary concludes the chapter and the dissertation.

REFERENCES

- Airasian, P. W. & Madaus, G. F. (1983). Linking testing and instruction: Policy issues. *Journal of Educational Measurement*, 20(2), 103-118.
- American Educational Research Association. (2000). *AERA position statement concerning high stakes testing in PreK-12 education*. Retrieved October 13, 2002, from <http://www.aera.net/about/policy/stakes.htm>.
- American Federation of Teachers (AFT), AFL-CIO. (2001). *Making standards matter*. Retrieved November 24, 2003 from <http://www.aft.org/>.
- Amrein, A. L., & Berliner, D. C. (2002). High-stakes testing, uncertainty, and student learning. *Education Policy Analysis Archives*, 10(18). Retrieved December 3, 2002 from <http://epaa.asu.edu/epaa/v10n18/>.
- Applebee, A. N., Langer, J. A. & Mullis, I. V. S. (1987). *The nation's report card: Learning to be literate in America: Reading*. Princeton, NJ: Educational Testing Service.
- Berliner, D. C. & Biddle, B. J. (1996). Making molehills out of molehills: Reply to Lawrence Stedman's review of *The Manufactured Crisis*. *Education Policy Analysis Archives* (4) 3. Retrieved November 10, 2003 from <http://epaa.asu.edu/epaa/v4n3.html>.
- Bidwell, C. E. (1965). The school as a formal organization. In J. March (Ed.), *Handbook of organizations* (pp. 974-1022). Chicago: Rand McNally.
- Blank, R.K. & Archbald, D. A. (1992). Manget schools and issues of education quality. *The Clearing House*. 66(2), 81-87.
- Bolger, N. & Kellaghan, T. (1990). Method of measurement and gender differences in scholastic achievement. *Journal of Educational Measurement*, 27, 165-174.
- Bond, L. A. & King, D. (1995). *State high school graduation testing: Status and recommendations*. Oakbrook, IL: North Central Regional Educational Laboratory.
- Borko, H. & Elliott, R. (1999). Hands-on pedagogy versus hands-off accountability: Tensions between competing commitments for exemplary math teachers in Kentucky. *Phi Delta Kappan*, 80, 394-400.
- Bowman, A. O., & Kearney, R. C. (1986). *The resurgence of the states*. Upper Saddle River, NJ: Prentice Hall.
- Boyd, W. L., & Kerchner, C. T. (1988). Introduction and overview: Education and the politics of excellence and choice. In W. L. Boyd & C. T. Kerchner (Eds.), *The*

- politics of excellence and choice in education* (pp.1-11). London: Taylor & Francis.
- Breckenridge, J. S. & Goldstein, D. (1998). A case study of Louisiana's SSI (LaSIP), In A. A. Zucker & P. M. Shields (Eds.), *SSI case studies, Cohort I: Connecticut, Delaware, Louisiana, and Montana*. Menlo Park, CA: SRI International.
- California Coalition for Authentic Reform in Education. (2002). *College and University Faculty Association of the National Council for the Social Studies*. Retrieved April 6, 2003 from <http://www.calcare.org/oldsite/resolutions/NCSS.htm>.
- Calkins, L., Montgomery, K., & Santman, D. (1998). *A teacher's guide to standardized reading tests: Knowledge is power*. Portsmouth, NH: Heinemann.
- Caporrimo, R. (Ed). (2001). Student perceptions, beliefs, and attitudes. *Academic Exchange Quarterly*, 5,2.
- Cardenas, J. A. (1994). Political limits to an education of value: The role of the state. In J. I. Goodlad, & P. Keating (Eds.), *Access to knowledge: The continuing agenda for our nation's schools* (pp. 273-286). New York: The College Board.
- Carnoy, M. & Loeb, S. (2002). Does external accountability affect student outcomes? A cross-state analysis. *Educational Evaluation and Policy Analysis*, 24(4) 305-331.
- Center for the Study of Testing, Evaluation, and Educational Policy. (1992). *The influence of testing on teaching math and science in grades 4-12* (Vols. 1-5). Boston College, Center for the Study of Testing, Evaluation, and Educational Policy.
- Chachkin, N. J. (1989). Testing in elementary and secondary schools. Can miscue be avoided? In B. Gifford (Ed.), *Test policy and the politics of opportunity allocation: The workplace and the law* (pp.163-187). Boston: Kluwer Academic Publishers.
- Clotfelter, C. T., & Ladd, H. F. (1996). Recognizing and rewarding success in public schools. In H. F. Ladd (Ed.), *Holding schools accountable: Performance-based reform in education* (pp. 23-63). Washington, DC: The Brookings Institution.
- Cohen, J. (1988). *Statistical power for the behavioral sciences*. New York: Academic Press.
- Commission on Secondary and Middle Schools (nd). Checklist of standards for the accreditation of middle schools. Retrieved September 15, 2004 from www.sacs.org/pub/elem/standards/standards.html.
- Commission on the Skills of the American Workforce (1990). *America's choice: High*

- skills or low wages!* Rochester, NY: National Center on Education and the Economy.
- Cook-Sather, Alison. (2002). Authorizing students' perspectives: Toward trust, dialogue, and change in education. *Educational Researcher*, 31, 3-14.
- Corbett, H. D. & Wilson, B. L. (1991). *Testing reform and rebellion*. Norwood, NJ: Ablex Publishing.
- Creswell, J. W. (1998). *Qualitative inquiry and research design: Choosing among five traditions*. Thousand Oaks, CA: Sage Publications.
- Danielson, M. L. (1999). How principals perceive and respond to a high stakes accountability measure. *Dissertation Abstracts International* 61, 03A. (UMI No. 9967692)
- Darling-Hammond, L. (1992). The future of teaching. *Educational Leadership*, 46, 4-10.
- Debra P. v. Turlington, 474 F. Supp. 244 (M.D. Fla. 1979); aff'd in part, rev'd in part, 644 F.2d 397 (5th Cir. 1981).
- Donato, R. & Lazerson, M. (2000). New directions in American educational history: Problems and prospects. *Educational Researcher*, 29(8), 4-15.
- Dorn, S. (1998). The political legacy of school accountability systems. *Education Policy Analysis Archives*, 6 (1). Retrieved September 10, 2003 from <http://epaa/asu.edu/epaa/v6n1.html>.
- Education Reporter. (1999, July). *NAEP Test Results Questioned*. Retrieved November 19, 2003 from <http://www.eagleforum.org/educate/1999/july99/naep.html>.
- Education Week. (2001, November 1). *Quality counts 2001*. Bethesda, MD.
- Education Week on the Web. (2004, January 9). Retrieved January 10, 2004 from <http://www.edweek.com/context/topics/issuespage.dfm?id=49>
- Educational Testing Service. (1988). *A summary of data collected from Graduate Record Examinations' test-takers during 1986-1987*. Data summary report No.12. Newark, NJ: Author.
- Elazar, D.J. (1984). *American federalism* (3rd ed.). New York: Harper & Row.
- Elementary and Secondary Education Act of 1965, Section 101.

- Ellwein, M. C., Glass, G. V. (1989). Ending social promotion in Waterford: Appearances and reality. In L. A. Shepard & M.L. Smith (Eds.), *Flunking grades: Research and policies on retention* (pp. 151-173). London: Falmer Press.
- Elmore, R. F., Abelman, C. H., & Fuhrman, S. H. (1996). The new accountability in state education reform: From process to performance. In H. F. Ladd (Ed.), *Holding schools accountable: Performance based reform in education* (pp. 65-98). Washington, DC: The Brookings Institution.
- Erickson, F., & Shultz, J. (1992). Students' experience of curriculum. In P.W. Jackson (Ed.), *Handbook of research on curriculum*. New York: Macmillan.
- Finley, S. J. (1999). The progress of education in Louisiana. *Southwest Educational Development Laboratory*. Retrieved October 15, 2002, from <http://www.sedl.org/pitl/pic/states/la/welcome.html>
- Finnigan, K. S., & Gross, B. M. (2001, April). *Teacher motivation and the Chicago probation policy*. Paper presented at the annual meeting of the Educational Research Association, Seattle, WA.
- Firestone, W. A. (1989). Using reform: Conceptualizing district initiative. *Educational Evaluation and Policy Analysis*, 11(2), 151-64.
- Firestone, W. A., Mayrowetz, D., & Fairman, J. (1998). Performance-based assessment and instructional change: The effects of testing in Maine and Maryland. *Educational Evaluation and Policy Analysis*, 20(2), 95-113.
- Fowler, F. C. (2000). *Policy studies for educational leaders*. Upper Saddle River, NJ: Merrill.
- Franklin, B. J., Pernici, S. C., & Yuan, X. S. (2001). *Grade level retention rates in Louisiana public schools: 1997-98 to 2000-01*. (ERIC Document Reproduction Service No. ED455301)
- Gall, M. D., Borg, W. R., & Gall, J. P. (2003). *Educational research: An introduction*. White Plains, NY: Longman.
- Gamoran, A. (1996). High-quality instruction for all. In A. T. Lockwood (Ed.). *Tracking: Conflicts and resolutions* (pp. 10-20). Thousand Oaks, CA: Corwin Press.
- Gampert, R. A. (1987). *A follow-up study of the 1982-83 promotional gates students*. New York: New York City Public Schools, Office of Educational Assessment.
- Garcia, G. E. & Pearson, P. D. (1993). Assessment and diversity. *Review of Research in Education*, 20, 337-391.

- Garcia, G. E., Stephens, D. L., Koenke, K. R., Pearson, P. D., Harris, V. J., & Jimenez, R. T. (1989). *A study of classroom practice related to the reading of low-achieving students: Phase one (Study 2.2.3.5)*. Urbana: University of Illinois; Reading Research and Educational Center.
- Gladfelter, H. R. (2000, January 10). Judge upholds Texas test against bias claims. *Education Daily*, pp. 1.
- Gleick, J. (1987). *Chaos: Making a new science*. New York: Penguin.
- Goertz, M., Floden, R., & O'Day, J. (1995). *Studies of education reform: Systemic reform Volume I: Findings and conclusions*. New Brunswick, NJ: Rutgers University, Center for Policy Research in Education.
- Goldring, E. & Smrekar, C. (2002). Magnet schools: Reform and race in urban education. *The Clearing House* 76(1), 12-15.
- Goodlad, J. I., (1984). *A place called school: Prospects for the future*. New York: McGraw-Hill Book Company.
- Gordon, S. P., & Reese, M. (1997). High stakes testing: Worth the price? *Journal of School Leadership*, 7, 345-368.
- Gottfredson, D. C., Fink, C. M., & Graham, N. (1994). Grade retention and problem behavior. *American Educational Research Journal*, 31(4), 761-784.
- Gordon, E. W. & Bonilla-Bowman, C. (1994) Equity and social justice in educational achievement. In R. Berne & L. O. Picus (Eds.) *Outcome equity in education* (pp.24-44). Thousand Oaks, CA: Corwin.
- Graham, P. (1980). Historians as policy makers, *Educational Researcher*, 9(11), 21-24.
- Grant, S. G. (2001). An uncertain lever: Exploring the influence of state-level testing on teaching social studies. *Teachers College Record*, 103(3), 398-426.
- Griffiths, D., Hart, A. H., & Blair, B. G. (1991). Still another approach to administration: Chaos theory. *Educational Administration Quarterly*, 27(3), 430-451.
- Hamilton, L. S. & Koretz, D.M. (2002). Tests and their use in test-based accountability systems. In L.S. Hamilton, B.M. Stecher, & S.P. Klein, (Eds.) *Making sense of test-based accountability in education* (pp. 13-49). Santa Monica, CA: RAND.
- Hamilton, L. S. & Stecher, B. M. (2002). Improving test-based accountability. In L. S. Hamilton, B. M. Stecher & S. P. Klein (Eds.) *Making sense of test-based accountability in education*. Santa Monica, CA: RAND.

- Hand, D. J. & Taylor, 1987). *Multivariate analysis of variance and repeated measures*. London: Chapman and Hall.
- Hauser, R. M., Pager, D. I. & Simmons, S. J. (2000, August). *Race-ethnicity, social background, and grade retention*. Presented at the annual meeting of the American Sociological Association, Washington, D.C.
- Herman, J. L. & Golan, S. (1993). The effects of standardized testing on teaching and Schools. *Educational Measurement, Issues and Practice*, 12(4), 20-25, 41-42.
- Heubert, J. P., & Hauser, R. M. (Eds.). (1999). *High stakes: Testing for tracking, promotion, and graduation*. National Research Council, Washington, DC: National Academy Press.
- Hochschild, H., & Scott, B., (1998). Trends: Governance and reform of public education in the United States. *Public Opinion Quarterly*, 62(1), 79-120.
- Hoffman, J. V., Assaf, L. C. & Paris, S. G. (2001) High-stakes testing in reading: Today in Texas, tomorrow? *Reading Teacher*, 54(5), 482-92.
- Hoffman, K. Llagas, C., Snyder, T. Status and trends in the education of Blacks. (2003, September). NCES Electronic Catalog Retrieved November 1, 2003 from <http://nces.ed.gov/pubsearch/pubsinfor.asp?pubid=2003034>
- Hopkins, K.D., & Glass, G. V. & Hopkins, B. R. (1987). *Basic statistics for the social sciences*. Englewood, Cliffs, NJ: Prentice Hall.
- Horn, C. Ramos, M., Blumer, I., Madaus, G. (2000). *Cut scores: Results may vary*. U. S. Department of Education, Office of Educational Research and Improvement, The National Board on Educational Testing and Public Policy Vol.1 No.1. (ERIC Document Reproduction Service No. 456 140).
- Jaeger, R. M. (1982). The final hurdle: Minimum competency achievement testing. In G. R. Austin & H. Garber (Eds.), *The rise and fall of national test scores*, (pp. 223-246). New York: Academic Press.
- Jaeger, R. M. (1989). Certification of student competence. In R. L. Linn (Ed.) *Educational Measurement*, 3rd Ed., (pp. 485-514). New York: Macmillan.
- Jennings, J. F. (1991). Chapter I: A view from Congress. *Educational Evaluation and Policy Analysis*, 13, 335.
- Jennings, N., & Spillane, J. (1996). State reform and local capacity: Encouraging ambitious instruction for all and local decision-making. *Journal of Educational Policy*, 11(4), 465-482.

- Johnson, C. S. (1942). The Negro public schools. In C. Washburne (Ed.) *Louisiana looks at its schools: A summary report of the Louisiana Educational Survey*. Baton Rouge: LA Louisiana Educational Survey Commission.
- Johnson, J. & Immerwahr, J., (1994). First things first: What Americans expect from the public schools. *New York Public Agenda*.
- Johnson, D. D. & Johnson, B. (2002). *High stakes: Children, testing, and failure in American schools*. Lanham, MD: Rowman and Littlefield.
- Jones, B. D. & Johnston, A.F. (2002, April). *The effects of high stakes testing on instructional practices*. Paper presented at the meeting of the American Educational Research Association, New Orleans, LA.
- Jones, M. G., Jones, B. D., Hardin, B. H., Chapman, L., Yarbrough, T., & Davis, M. (1999). The impact of high stakes testing on teachers and students in North Carolina. *Phi Delta Kappan*, 81(3), 193-203.
- Jones, M. G., Jones, B. D., & Hargrove, T. Y. (2003). *The unintended consequences of high stakes testing*. Lanham, MD: Rowman & Littlefield.
- Jourard, S. M. (1964). *The transparent self*. Princeton, HJ: Van Nostrand.
- Kaplan, M., & O' Brien, S. (1991). *The governors and the new federalism*. Boulder, CO: Westview Press.
- Kelley, C., Odden, A., Milanowksi, A., & Heneman, H. (2000). *The motivational effects of school-based performance awards* (CPRE Policy Brief No. RB-29). Philadelphia, PA: Consortium for Policy Research in Education.
- Kiely, K. & Henry, T. (2001, December 17). Will No Child Be Left Behind? *USA Today*, pp. D4.
- Kober, N. (1992). *The role and impact of Chapter I, ESEA, evaluation and assessment practices*. Washington, DC: U. S. Government Printing Office, Chapter I.
- Kohn, A. (2000, September 27). Standardized testing and its victims. *Education Week*, 20(4), 46-47, 60.
- Koretz, D. (1988). Arriving at Lake Wobegon: Are standardized tests exaggerating achievement and distorting instruction? *American Educator*, 12 (2) 3-15, 46-52.
- Koretz, D. (1992). What happened to test scores, and why? *Educational Measurement: Issues and Practice*, 11, 7-11.

- Koretz, D. & Diebert, E. (1993). *Interpretations of National Assessment of Educational Progress (NAEP) anchor points and achievement levels by the print media in 1991*. Santa Monica, CA: The RAND Corporation. (ERIC Document Reproduction Service No. ED 367683)
- Krueger, R. A. (1988). *Focus groups: A practical guide for applied research*. Newbury Park, CA: Sage.
- Krueger, R. A. (1994). *Focus groups: A practical guide for applied research*. Thousand Oaks CA: Sage.
- Krueger, R. A. (2004 a). *Asking questions that yield powerful information*. Retrieved March 12, 2004 from University of Minnesota Web site:
http://www.tc.umn.edu/~rkrueger/focus_aq.html
- Krueger, R. A. (2004b) *Systematic analysis process*. Retrieved March 12, 2004 from the University of Minnesota Website:
http://www.tc.umn.edu/~rkrueger/focus_analysis.html
- Krueger, R. A. & Casey, M.A. (2000). *Focus groups. A practical guide for applied research*. 3rd ed. US: Sage.
- Ladd, H. F. (1999). The Dallas school accountability and incentive program: An evaluation of its impacts on student outcomes. *Economics of Education Review*, 18, 1-16.
- LaSIP Overview. Retrieved, February 10, 1997 from <http://www.lasip.state.la.us> currently unavailable.
- Le, V. & Klein, S. P. (2002). Technical criteria for evaluating tests. In L. S. Hamilton, B. M. Stecher, & S. P. Klein (Eds.) *Making sense of test-based accountability in education*. Santa Monica, CA: RAND.
- Lewis, A. (2000). *High stakes testing: Trends and issues*. Mid-continent Research for Education and Learning. Retrieved November 24, 2003, from
<http://www.mcrel.org/pdfconversion/policybriefs/pb%5Fhighstakestesting.html>.
- Linn, R. L. (2000). Assessments and accountability. *Educational Researcher*, 29(2), 4-16.
- Linn, R. L., Graue, M. E., Sanders, N. M. (1990). *Comparing state and district test results to national norms: Interpretations of scoring "above the national average"*. CSE report no. 308). Center for Research on Evaluation, Standards, and Student Testing. U.S. Department of Education, Office of Educational Research and Improvement, Educational Resources Information Center.

- Louisiana Department of Education . (2001, 2003). *LEAP for the 21st century interpretive guide: Grades 4,8, 10 and 11 criterion-referenced tests*. Retrieved April 15, 2004 and September 15, 2004 from <http://www.louisianaschools.net/lde/uploads/1278.pdf>.
- Louisiana Department of Education. (1999). *Louisiana state education progress report 1999-2000*. Retrieved October 19, 2003, from <http://www.doe.state.la.us/lde/pair/StateReport9900/210080TxtA.pdf>.
- Louisiana Department of Education. (2000a February). Reaching for results LEAP 21. Retrieved December 15, 2002, from <http://www.doe.state.la.us/DOE/PDFsintro>
- Louisiana Department of Education. (2000b, June) *State regulation of private schools*. Retrieved November 25, 2003, from <http://www.ed.gov/pubs/RegPrivSchl/louisiana.html>
- Louisiana Department of Education. (2001a). *Pre-GED/Skills option program resource guide*. Baton Rouge, LA
- Louisiana Department of Education. (2001b). *What is LEAP 21/GEE 21?* Retrieved October 30, 2003, from <http://www.doe.state.la.us/lde/uploads/1703.pdf>.
- Louisiana Department of Education. (2003). LEAP for the 21st century high stakes testing policy. Retrieved January 1, 2004, from <http://www.doe.state.la.us/lde/uploads/1603.pdf>
- Louisiana Department of Education. (2004). *Student standards and assessments: Grade-level expectations*. Retrieved June 17, 2004 from <http://www.louisianaschools.net/lde/ssa/1915.html>
- Louisiana LEARN Commission, (1996). *LEARN for the 21st century: comprehensive plan to improve education*. Baton Rouge, LA
- Madaus, G. F. 1988. The distortion of teaching and testing: high stakes testing and instruction. *Peabody Journal of Education* 65, 29-46.
- Madaus, G. F. 1991. The effects of important tests on students. *Phi Delta Kappan*, 73, 226-231.
- Madaus, G. F., & Tan, A. G. (1993). The growth of assessment. In G. Cawelti (Ed.), *Challenges and achievements of American education: The 1993 ASCD yearbook* (pp. 53-79). Alexandria, VA: Association for Supervision and Curriculum Development.

- Marshall, C. & Rossman, G. B. (1999). *Designing qualitative research* (3rd ed.). Thousand Oaks: Sage.
- Maxcy, S. J. (1995). *Democracy, chaos, and the new school order*. Thousand Oaks, CA: Corwin Press.
- Mazzeo, J., Schmitt, A. P., & Bleistein, C. A. (1993). *Sex-related performance differences on constructed response and multiple-choice sections of advanced placement examinations* (CB Report No. 92-ETS RR-93-5). New York: College Entrance Examination Board.
- Mazzoni, T. L., (1995). State policy-making and school reform: Influences and influentials. In J. D. Scribner & D. H. Layton (Eds.), *The study of educational politics* (pp. 53-73). London: Falmer Press.
- McDonnell, L. M. (2002). Accountability as seen through a political lens. In L.S. Hamilton, B.M. Stecher, & S. P. Klein, (Eds.) *Making sense of test-based accountability in education*. Santa Monica, CA: RAND.
- McNeil, L. M. (2000). Creating new inequalities: Contradictions of reform. *Phi Delta Kappan*, 81, 729-734.
- Mehrens, W. A. (1998). Consequences of assessment: What is the evidence? *Vice presidential address for division D, annual meeting of the American Educational Research Association*, San Diego, CA.
- Meier, D. (2002). Standardization versus standards. *Phi Delta Kappan*, 84(3), 190-198.
- Meyer, R. H. (1996). Comments on chapters two, three, and four. In H. F. Ladd (Ed.), *Holding schools accountable: Performance-based reform in education* (pp.137-145). Washington, DC: The Brookings Institution.
- Morgan, D.L. (1997). *Focus groups as qualitative research* (2nd ed.). Thousand Oaks, CA: Sage.
- Mullis, I. V. S. & Jenkins, L. B. (1990). The reading report card, 1971-88. *Trends from the nation's report card*. Princeton, NJ: National Assessment of Educational Progress, Educational Testing Service.
- Nathan, R. P. (1993). The role of the states in American federalism. In C. E. Van Horn (Ed.), *The state of states* (2nd ed.), (pp. 15-32). Washington, DC: Congressional Quarterly Press.

- National Catholic Education Association. (May, 2004). *National Catholic Educational Association statement on accountability and assessment in Catholic education*. Retrieved October 23, 2004 from <http://www.ncea.org/publicpolicy/policystatements/accountability.asp>
- National Center for Educational Statistics. (1988). *Education indicators*. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement.
- National Center for Education Statistics. (1996). Remedial Education at Higher Institutions in Fall 1995 (Washington, D.C.: U.S. Department of Education, 1996), p.34.
- National Center for Education Statistics (2002, February). *School locale codes 1987-2000*. NCES 200202. Retrieved July 26, 2004 from <http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=200202>
- National Center for Education Statistics. (2003, May). *Overview of public elementary and secondary schools and districts: School year 2001-02*. NCES 2003-411. Retrieved November 3, 2003, from: http://nces.ed.gov/pubs2003/overview03/table_11.asp.
- National Center for Education Statistics. (nd). Information on Public Schools and School Districts in the United States. Retrieved January 6, 2004 from: <http://nces.ed.gov/ccd/aboutCCD.asp>
- National Council of Teachers of English. (2000). *On urging reconsideration of high stakes testing*. Retrieved November 25, 2003 from <http://www.ncte.org/about/over/positions/category/assess/107357.htm>.
- National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. [Electronic version]. Available at <http://www.ed.gov/pubs/NatAtRisk/intro.html>
- National Council on Education Standards and Testing (1992). *Raising standards for American education: A report to Congress, the Secretary of Education, the National Goals Panel, and the American people*. Washington, DC: U.S. Government Printing Office.
- National PTA, (1996). *High stakes testing: Is it fair to students?* Retrieved October 26, 2002 from http://www.pta.org/parentinvolvement/helpchild/oc_highstake2.asp
- Natriello, G. & McDill, E. L. (1999). Title I: From funding mechanism to educational program. In The Harvard Civil Rights Project, *Hard work for good schools: facts not fads in Title I reform*. Cambridge, MA: Harvard University.

- Newton, R. R. & Rudestam, K. E. (1999). *Your statistical consultant: Answers to your data analysis questions*. Thousand Oaks, CA: Sage Publications.
- No Child Left Behind Act of 2001, Pub. L., No. 107-110, 115 Stat. 1425 (2002).
- O'Day, J. A. (2002). Complexity, accountability, and school improvement. *Harvard Educational Review*, 72 (3), 293-329.
- O'Day, J. A. & Smith, M. S. (1993). School reform and equal opportunity: An introduction to the education symposium. *Stanford Law and Policy Review*, 4, 15-20.
- Office of Technology Assessment. (1992). *Testing in American schools: Asking the right questions*, (OTA-SET-519), Washington, DC: U. S. Government Printing Office.
- Olson, C. L. (1976). On choosing a test statistic in multivariate analysis of variance. *Psychological Bulletin*, 83(4), 579-586.
- Phelan, P.L., Davidson, A. L., & Cao, H. T. (1992). Speaking up: Students' perspectives on school. *Phi Delta Kappan*, 73(9), 695-704.
- Popham, W. J. 2000. *Modern educational measurement: Practical guidelines for educational leaders* (3rded.). Needham, MA: Allyn and Bacon.
- Porter, A. C., Archbald, D. A. & Tyree, A. K. Jr. (1991). Reforming the curriculum: Will empowerment policies replace control? In S. H. Fuhrman & B. Malen (Eds.), *The politics of curriculum and testing* (pp. 11-36). London: Falmer Press.
- Quality Counts. (2003). *To close the quality gap, quality counts*. 22 (17), p. 7.
- Ravitch, D. (1995). *Debating the future of American education: Do we need national standards and assessments?* Washington, DC: The Brookings Institution.
- Ravitch, D. (2000). *Left back. A century of failed school reform*. New York: Simon and Schuster.
- Ravitch, D., & Vinovskis, M. (1995). *Learning from the past: What history teaches us about reform*. Baltimore: Johns Hopkins University Press.
- Reardon, S. F., (1996, April). *Eighth grade minimum competency testing and early high school dropout patterns*. Paper presented at the annual meeting of the American Educational Research Association, New York.

- Rebell, M. A. (1989). Testing public policy and the courts. In B. Gifford (Ed.) *Test policy and the politics of opportunity allocation: The workplace and the law* (pp. 135-162). Boston: Kluwer Academic Publishers.
- Resnick, D. P. (1981). Testing in America: A supportive environment. *Phi Delta Kappan*, 62(9), 625-628.
- Robelen, E. W., (2000, May 24). La. set to retain 4th, 8th graders based on state exams. *Education Week*, 19(37) 21-24.
- Roderick, M., Jacob, B. A., & Bryk, A. S. (2002). The impact of high stakes testing in Chicago on student achievement in promotional gate grades. *Educational Evaluation and Policy Analysis*, 24(4), 333-358.
- Rose, L. C. & Gallup, A. M. (2004).. The 36th annual Phi Delta Kappa/Gallup poll of the public's attitudes toward the public schools. Retrieved September 9, 2004 from: <http://www.pdkintl.org/kappan/k0409pol.htm>
- Rothman, R. (1992, October 21). Study confirms fears regarding commercial tests. *Education Week*, pp. 1, 13.
- Rothman, R. (1996). Taking aim at testing. *American School Board Journal*, 183,(2), 27-30.
- Rothman, R., Slattery, J. B. & Vranek, J.L. (Achieve, Inc.) & Resnick, L. B. (2002). *Benchmarking and Alignment of Standards and Testing* (CSE Technical Report 566). CA: University of California, Los Angeles, Graduate School of Education & Information Studies.
- San Miguel, G. (1987). *"Let all of them take heed" Mexican Americans and the campaign for educational equality in Texas, 1910-1981*. Austin University of Texas Press.
- Shepard, L. A. (1990). Inflated test score gains: Is the problem old norms or teaching to the test? *Educational Measurement: Issues and Practice*, 9(3), 15-22.
- Shepard, L. A. 1991. Will national tests improve student learning? *Phi Delta Kappan*, 73, 232-238.
- Shippo, D. & Firestone, W. (2003, June 18). Juggling accountabilities. *Education Week*, 22(41), pp. 45,56.
- Simon, H. (1965). *Administrative behavior: A study of the decision-making processes in administrative organization* (2nd ed.). New York: Free Press. (Original work published 1947).

- Smith, M. L. (1991). Put to the test: The effects of external testing on teachers. *Educational Researcher*, 20, 8-11.
- Smith, M. L., Miller-Kahn, L., Heineche, W., & Jarvis, P.F. (2004). *Political spectacle: And the fate of American schools*. New York: RoutledgeFalmer.
- Smith, M. L., & Rottenberg, C. 1991. Unintended consequences of external testing in elementary schools. *Educational Measurement: Issues and Practice*, 10, 7-11.
- Smith, M. S. & Jenkins, J. W. (1982). Legislation. In *Encyclopedia of educational research*, 5th Edition, H. F. Mitzel, (Ed.). New York: The Free Press.
- Smith, M. S. & O'Day, J. A. (1990). Systemic school reform. In M. E. Goetz & D. E. Mitchell (Eds.), *Politics of education association yearbook* (pp. 233-267). London: Taylor & Francis.
- Smith, M. S., O'Day, J. A., & Cohen, D. K. (1990). National curriculum American style: What might it look like? *American Educator*, 14(4), 10-17, 40-43.
- Soloman, P. (1998). *The curriculum bridge: From standards to actual classroom practice*. Thousand Oaks, CA: Corwin.
- Spillane, J. (1998). The progress of standards-based reforms and the non-monolithic nature of the local school district: Organizational and professional considerations. *American Educational Research Journal*, 35(1) 33-63.
- Spillane, J. (2000). Cognition and policy implementation: District policy-makers and the reform of mathematics education. *Cognition and Instruction*, 18, 141-179.
- Spillane, J. (2002). Local theories of teacher change: The pedagogy of district policies and programs. *Teachers College Record*, 104, 377-420.
- SPSS Base 8.0 User's Guide (1998). Chicago: SPSS Inc.
- Stake, R. E. (1995). *The art of case study research*. Thousand Oaks, CA: Sage Publications.
- Stedman, L. C. (1996). The achievement crisis is real: A review of the manufactured crisis [Electronic version]. *Education Policy Analysis Archives* (4)1. Retrieved November 1, 2003 from <http://epaa.asu.edu/epaa/v4n1.html>
- Stroufe, G. E. (1995). Politics of education at the federal level. In J. D. Scribner & D. H. Layton (Eds.) *The study of educational politics*. Washington DC: The Falmer Press.

- Swanson, C. B. & Stevenson, D. L. (2002). Standards-based reform in practice: Evidence on state policy and classroom instruction from the NAEP state assessments. *Educational Evaluation and Policy Analysis*, 24(1), 1-27.
- Tabachnick, B. G. & Fidell, L. S. (2001). *Using multivariate statistics*, 4th ed. Boston: Allyn & Bacon.
- Tashakkori, A. & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches*. Thousand Oaks, CA: Sage.
- Thernstrom, S. and Thernstrom, A. (1997). *America in Black and White: One nation, indivisible*. New York: Simon & Schuster.
- Thompson, B. (1994). The concept of statistical significance testing. *Practical Assessment, Research and Evaluation*, 4(5). Retrieved October 9, 2004 from <http://PAREonlin.net/getvn.asp?v=4&n=5>
- Timar, T. (1994). Program design and assessment strategies in Chapter I. In K. C. Wong & M. C. Wang (Eds.), *Rethinking Policy for At-Risk Students* (p.67). Berkeley, CA: McCutchan Publishing.
- Tye, K. A. (1992). Restructuring our schools: Beyond the rhetoric. *Phi Delta Kappan*, 74(1), 8-14.
- U.S. Census Bureau. (2002) *Ranking table: 2002 median household income*. Retrieved December 10, 2003, from <http://www.census.gov/acs/www/Products/Ranking/2002/R07T040.htm>.
- U.S. Census Bureau. (May 28, 2003a). *Profile of general demographic characteristics: Census 2000 summary file 1 (SF1)*. Retrieved November 25, 2003, from http://factfinder.census.gov/servlet/QTTTable?ds_name=DEC_2000_SF1_U&geo_id=04000US22&qr_name=DEC_2000_SF1_U_DP1.
- U.S. Census Bureau. (May 28, 2003b) *Profile of selected economic factors: census 2000 summary file 3 (SF3)*. Retrieved November 25, 2003 from http://factfinder.census.gov/servlet/QTTTable?ds_name=DEC_2000_SF3_U&geo_id=04000US22&qr_name=DEC_2000_SF3_U_DP3.
- United States Department of Education. (1995). *Progress of education in the United State of America: 1990-1994*. Retrieved November 24, 2003, from <http://www.ed.gov/pubs/Prog95>.
- United States Department of Education. (1998). *School poverty and academic performance: NAEP achievement in high-poverty schools*. Retrieved November 25, 2003, from <http://www.ed.gov/pubs/schoolpoverty/index.html>.

- United States Department of Education (2002). *Vocational education offerings in rural high schools* (NCES 2002-120). Washington, DC: Hudson, L & Shafer, L.
- Virovskis, M. (1999). *History and educational policymaking*. New Haven: Yale University Press.
- Walker, V. S. (1996). *Their highest potential: An African American school community in the segregated south*. Chapel Hill: University of North Carolina Press.
- Wallace, B. & Graves, W. (1995). *Poisoned Apple*. New York: St. Martin's Press.
- Washor, E. & Mojkowski, C. (2003, April). Accountability in small schools. *Education Week*, 22(30), 40, 42.
- Wayne, A. (2002, June 13). Teacher inequality: New evidence on disparities in teachers' academic skills. *Education Policy Analysis Archives*, 10(30). Retrieved August 2, 2002 from <http://epaa.asu.edu/epaa/v10n30/>.
- Weick, K. (1976). Educational organizations as loosely coupled systems. *Administrative Science Quarterly*, 2(1), 1-19.
- Wideen, M. F., O'Shea, T., Pye, I., & Ivany, G. (1997). High stakes testing and the teaching of science. *Canadian Journal of Education*, 22(4), 438-444.
- Willingham, W. W., & Cole, N. S. (1997). *Gender and fair assessment*. Mahwah, NJ: Erlbaum.
- Wolf, S. A., & Borko, H., McIver, M. C., & Elliott, R. (1999). *No excuses: School reform efforts in exemplary schools of Kentucky*. (CSE Technical Report 514). Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing.
- Wolf, S. A., & McIver, M. C. (1999). When process becomes policy. *Phi Delta Kappan*, 80(5), 401-406.
- Yin, R. K. (1994). *Case study research: Design and methods*. (2nd ed.) Newbury Park, CA: Corwin Press.

APPENDIX A
LETTERS – SUPERINTENDENT, PRINCIPAL, AND PARENT

January 30, 2004

School Superintendent:
School District Anywhere
City, State Zip Code

Dear School Superintendent:

I am requesting permission to visit in high school(s) in your parish to collect data for a research study. I am conducting the study as a part of my doctoral dissertation at Louisiana State University. As a purpose for the study, I will examine differences in high stakes performance between Black and White students across different school community types. In addition, parent perspectives of high stakes testing impacts will be explored.

My research includes all safeguards as established by LSU's Institutional Review Board. The safeguards that I employ include confidentiality in all data collected.

The knowledge that I gain can mean increased benefits to the students in your parish. If you have any questions, please feel free to contact me at 985-384-0861 or e-mail me: momancuso@atvci.net. Please indicate your permission for this data collection to take place.

Sincerely,

Monica L. Mancuso

_____ Permission is granted for Mrs. Mancuso to conduct a study in _____ Parish Schools.

_____ (Signed) _____ (Date)

January 30, 2004

School Principal
Public School
Somewhere, LA

Dear School Principal:

I am requesting permission to visit your school to collect data for a research study. I am conducting the study as a part of my doctoral dissertation at Louisiana State University. As a purpose for my study, I will examine the achievement differences between Black and White students on high stakes tests across different school community types.

In order to collect the data, I want to conduct parent focus groups and interviews.

Your superintendent has given permission for me to pursue data collection with your approval. I feel that much can be learned from studying the high stakes test performances of Black and White students across different school community types.

I understand the pressures and time constraints within a school setting, and I will minimize the time I spend involved in data collection. My research includes all safeguards as established by LSU's Institutional Review Board. The safeguards that I will employ include confidentiality in all data collection.

The knowledge I gain can mean increased benefits to the students in your school. If you have any questions, please feel free to contact me at 985-384-0861 or e-mail me: momancuso@atvci.net.

Please indicate your permission for these observations to take place.

Sincerely,

Monica L. Mancuso, LSU

_____ Permission is granted for Mrs. Mancuso to conduct a study in _____ School.

_____ (Signed) _____ (Date)

Project Title: Effect of High Stakes Testing in Louisiana

Performance Site: Selected middle and high schools in the state

Research Investigator: Monica L. Mancuso, LSU (Questions may be addressed to
Monica L. Mancuso, M-F 8-4 P.M. 337-836-9661 Ext 3311 or
985-518-2009) You may also contact Dr. Dianne Taylor 1-225-
578-6900

Purpose: Our schools are interested in reducing the achievement differences in high stakes test performance. Educators work toward meeting annual yearly achievement goals across sub-groups of students in the accountability program in our state. Examining the impact of high stakes testing on students in Louisiana is the purpose for this study.

Inclusion Criteria: Parent perspectives of students in grades 9 and 8.5, who tested in the spring of 2003, are included.

Exclusion Criteria: Parents of non-eighth grade spring 2003 testers are excluded.

Description of the Study: Parent perspectives are needed as parents are often left out of discussions on mandated state testing. An outside research student from LSU will collect data through parent focus groups and interviews. The researcher will ask questions about how the high stakes test (LEAP 21) has impacted students taking the test. Randomly selected parents are asked to participate in interviews. The focus group interview will last from 50-60 minutes. The results of the focus group interview will help educators understand high stakes test performance within their own school. Some parents are selected to participate in follow-up interviews, with the purpose of clarifying perspectives.

Benefits: Educators can decrease negative test impacts on students by utilizing parent perspectives to understand test impacts.

Risks: There are no known risks for participation.

Right to Refuse: Participation is voluntary. A parent may withdraw at any time without penalty or loss of any benefit to which they might be entitled.

Privacy: Results of the study may be published, but no names or identifying information will be used. Subject identity is confidential unless required by law.

Financial Information: There is no cost for participation in the study.

Signatures

The study has been explained to me. If I have questions, I may call the research investigator.

Parent Signature _____ School _____

Phone Number _____ Date _____

Best time to call _____

Best time to meet _____

APPENDIX B FOCUS GROUP QUESTIONS

Reform Intent

PROBES: Would you explain further?
Would you give an example?
I don't understand.

1. What made students have to take LEAP 21 to pass to ninth grade? (Introductory question) (IF THE RESPONSE IS THE STATE OF LOUISIANA, PROBE FOR WHAT REASONING WENT INTO THE STATE REQUIRING THIS TEST)
2. Tell me about whether LEAP 21 is making students learn more. (Key question – reform intent)

The Test and Student Learning

1. How has the LEAP 21 test affected what your child is taught at school? (Key question – Impact on student learning)
2. * What can you tell me about English/Language Arts and mathematics LEAP tests? Do students have a more difficult time with a particular subject”?
3. When a student passes the test, what do you think that says about the student’s learning? What about when a student fails? (Key question- student learning)

Test Fairness

1. What do you think about having to pass a test to be promoted? (Key question – single measure for promotion)
2. What is your opinion of taking the test to be promoted to ninth grade in the spring of the eighth grade year? (Key question Timing of the test)

3. Give your opinion of how the test impacts groups of students – for example rich or poor PAUSE FOR A RESPONSE, Black or White PAUSE FOR A RESPONSE, male or female PAUSE FOR A RESPONSE)? (Key question– impact on groups of students)
4. What is your opinion about non-public school students being promoted without passing a test? To what extent did you consider non-public school for your child? (Key question–test exemption for nonpublic students)
5. Of all the things that we have talked about, what do you feel is the most important? (Ending question)

Summarize

3 Step Conclusion: summarize, review purpose, and ask if anything missed- THANKS
AND DISMISSAL

* The question was added for the urban focus group sessions.

APPENDIX C
FORM FOR POST FOCUS GROUP NOTES

Seating Chart Sketch

Debriefing Remarks

Themes, Hunches, Interpretations, Ideas

Comparisons and contrasts to other focus groups

NOTE: LABEL AND FILE TAPES

VITA

Monica Laughlin Mancuso was born in Hobbs, New Mexico, on October 12, 1953. She is the daughter of Earl and Mary Catherine Laughlin of Morgan City, Louisiana. After graduating in 1971 from Central Catholic High School in Morgan City, Ms. Mancuso received a Bachelor of Science degree in horticulture from Louisiana State University in 1975 and a Bachelor of Arts degree with honors in elementary education from the University of Louisiana at Lafayette in 1982. Further studies earned Ms. Mancuso a Master of Education degree with honors from Nicholls State University in 1989.

Ms. Mancuso began her teaching career in 1975 and taught for twelve years before being named principal of Shannon Elementary, a fourth through sixth grade school in Morgan City. Ms. Mancuso served ten years as a school principal, including four years at Morgan City High School, before being named as supervisor for the St. Mary Parish School Board, a position that Ms. Mancuso currently holds.

Ms. Mancuso and her husband, Bart, have four children: Dr. Frances Chauvin, Mrs. Mary White, Andrew Mancuso and Luke Mancuso. Ms. Mancuso has two publications in the *Journal of Philosophy and History of Education*. Both of the publications were presented at the annual meeting of the society. The fifty-third volume of the journal contains the first publication, "John Dewey's Model for 'Figuring It Out': Utilizing a Deweyan Methodological Approach for Twenty-first Century Reform." The second publication is awaiting printing and is titled "Dewey's Term of 'Backwardness' Becomes a 'Forwardness': Student Perspectives in Louisiana's Skills Option Track."